GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 06:07:38 ; Search time 521.76 Seconds
(without alignments)
727.227 Million cell updates/sec

Title: US-09-846-456-4
Perfect score: 221
Sequence: 1 gtaattgcgagcgagagtga.....aacacaaaagtggaaaacag 221

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	221	100.0	7260	22	AAD21326	Human ATP binding
2	221	100.0	7260	22	AAI70315	Human ATP binding
3	219	99.1	763	22	AAH04729	Human cDNA clone (
4	219	99.1	1750	22	AAH17451	Human cDNA sequenc
5	217	98.2	736	22	AAH07432	Human cDNA clone (
6	217	98.2	1556	22	AAH18606	Human cDNA sequenc
7	205	92.8	7086	22	ABA09200	Human ABCA1 homolo
8	205	92.8	7086	22	AAK52667	Human polynucleoti
9	205	92.8	7281	22	AAK51683	Human polynucleoti
10	205	92.8	9854	22	AAS06121	Human ABC1 DNA seq
11	201	91.0	183999	22	AAF92831	Human ABC1 genomic
12	197	89.1	227	21	AAC09615	Human secreted pro
13	197	89.1	10442	22	AAF24680	Nucleotide sequenc
14	197	89.1	10442	22	AAF24702	Nucleotide sequenc
15	188	85.1	10474	22	AAF24685	Nucleotide sequenc
16	188	85.1	10474	22	AAF24686	Nucleotide sequenc
17	188	85.1	10474	22	AAF24707	Nucleotide sequenc
18	188	85.1	10474	22	AAF24708	Nucleotide sequenc
19	92	41.6	446	22	AAS04035	Partial human ABC1
20	92	41.6	9741	22	AAS06120	Human ABC1 DNA seq
21	91	41.2	1643	22	AAF24681	Nucleotide sequenc
22	91	41.2	1643	22	AAF24703	Nucleotide sequenc
c 23	19	8.6	310	21	AAC00305	Human secreted pro
24	19	8.6	447	20	AAV90187	EST clone DH318.
c 25	19	8.6	710	11	AAQ06308	Sequence of DNA fr
c 26	19	8.6	735	15	AAQ73229	Soluble human inte
c 27	19	8.6	2184	11	AAQ06301	Sequence encoding
c 28	18	8.1	367	21	AAC13639	Human secreted pro
c 29	18	8.1	710	22	AAF58344	Human GTP-binding
c 30	18	8.1	758	21	AAF15226	Trichoderma reesei
31	18	8.1	1166	23	AAS68374	DNA encoding novel
c 32	18	8.1	1287	23	AAS72684	DNA encoding novel
33	18	8.1	1656	15	AAQ66990	DF3 enhancer. Hom
34	18	8.1	1725	15	AAQ66988	5' flanking region
c 35	18	8.1	2484	22	ABA20286	Human nervous syst
c 36	18	8.1	2484	22	ABA20287	Human nervous syst
c 37	18	8.1	2779	21	AAZ86925	Netrin-2 coding se
c 38	18	8.1	2783	16	AAQ92367	Chick p75 cDNA. G
c 39	18	8.1	9248	20	AAZ32011	Human METH1 relate
c 40	18	8.1	9248	22	AAC90068	AB001735 cDNA clon
41	18	8.1	17294	24	ABL32987	Human immune syste
42	18	8.1	21436	22	AAK70011	Human immune/haema
43	18	8.1	21436	22	AAK79799	Human immune/haema
44	17	7.7	17	22	AAF92943	Wild type sequence
c 45	17	7.7	125	22	AAK90657	Human digestive sy

ALIGNMENTS

RESULT 1
AAD21326
ID AAD21326 standard; DNA; 7260 BP.
XX
AC AAD21326;
XX
DT 28-JAN-2002 (first entry)
XX
DE Human ATP binding cassette transporter 1 (ABC1) gene.
XX
KW Human; ATP binding cassette transporter 1; ABC1; coronary heart disease;
KW dermatological; atherosclerosis; cardiovascular; inflammatory disease;
KW psoriasis; lipid disorder; antibacterial; septic shock; gene therapy;
KW immunosuppressive; lupus erythematosus; rheumatoid arthritis; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 321..7106
FT /*tag= a
FT /product= "Human ABC1 protein"
XX
PN EP1136552-A1.
XX
PD 26-SEP-2001.
XX
PF 20-MAR-2000; 2000EP-0105820.
XX
PR 20-MAR-2000; 2000EP-0105820.
XX
PA (FARB) BAYER AG.
XX

```
PI   Schmitz G,  Bodzioch M;
XX
DR   WPI; 2001-640388/74.
DR   P-PSDB; AAE13022.
XX
PT   New adenosine triphosphate binding cassette transporter-1 gene
PT   polymorphisms, useful for diagnosing and treating lipid disorders,
PT   cardiovascular diseases and inflammatory diseases  -
XX
PS   Example 1; Fig 1; 48pp; English.
XX
CC   The invention relates to four common polymorphisms in the gene encoding
CC   ATP-binding cassette transporter-1 (ABC1). ABC1 is associated with
CC   decreased ApoA-1 mediated efflux of cholesterol. The polymorphisms in
CC   ABC1 directly affects cellular lipid homeostasis, which is a key factor
CC   in the atherogenetic processes. The ABC1 polymorphisms are useful for
CC   diagnosing and treating lipid disorders, cardiovascular diseases
CC   (coronary heart disease, atherosclerosis) and inflammatory diseases
CC   (psoriasis, lupus erythematosus). The identification of ABC1 as a
CC   transporter for interleukin-1beta (IL-1beta) identifies this gene as
CC   a candidate for treatment of inflammatory diseases including rheumatoid
CC   arthritis and septic shock. The present sequence is human ABC1 gene.
XX
SQ   Sequence 7260 BP; 1834 A; 1765 C; 1905 G; 1756 T; 0 other;


  Query Match             100.0%;  Score 221;  DB 22;  Length 7260;
  Best Local Similarity   100.0%;  Pred. No. 3e-103;
  Matches  221;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 gtaattgcgagcgagagtgagtggggccgggacccgcagagccgagccgacccttctctc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8 gtaattgcgagcgagagtgagtggggccgggacccgcagagccgagccgacccttctctc 67

Qy     61 ccgggctgcggcagggcagggcggggagctccgcgcaccaacagagccggttctcagggc 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     68 ccgggctgcggcagggcagggcggggagctccgcgcaccaacagagccggttctcagggc 127

Qy    121 gctttgctccttgtttttttccccggttctgttttctccccttctccggaaggcttgtcaa 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    128 gctttgctccttgtttttttccccggttctgttttctccccttctccggaaggcttgtcaa 187

Qy    181 ggggtaggagaaagagacgcaaacacaaaagtggaaaacag 221
          |||||||||||||||||||||||||||||||||||||||||
Db    188 ggggtaggagaaagagacgcaaacacaaaagtggaaaacag 228


RESULT   2
AAI70315
ID   AAI70315 standard; cDNA; 7260 BP.
XX
AC   AAI70315;
XX
DT   07-JAN-2002  (first entry)
XX
DE   Human ATP binding cassette transporter 1 (ABC1) cDNA.
XX
KW   ATP binding cassette transporter 1; ABC1; human; lipid disorder;
KW   cholesterol; cardiovascular disease; inflammatory disease;
KW   antiinflammatory; antilipaemic; antipsoriatic; dermatological;
KW   Tangier disease; coronary heart disease; diagnosis; gene therapy;
KW   polymorphism; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             321..7106
FT                   /*tag=  a
FT   CDS             501..7106
FT                   /*tag=  b
FT                   /note= "alternative open reading frame of AAI70314"
FT   variation       replace(976,A)
FT                   /*tag=  c
FT   variation       replace(1516,C)
FT                   /*tag=  d
FT   variation       replace(2969,G)
FT                   /*tag=  e
FT   variation       replace(3836,C)
FT                   /*tag=  f
XX
PN   EP1136554-A1.
XX
PD   26-SEP-2001.
XX
PF   24-MAR-2000; 2000EP-0106401.
XX
PR   24-MAR-2000; 2000EP-0106401.
XX
PA   (FARB ) BAYER AG.
XX
PI   Schmitz G,  Bodzioch M;
XX
DR   WPI; 2001-640389/74.
DR   P-PSDB; AAM50228.
XX
PT   New adenosine triphosphate binding cassette transporter gene
PT   polymorphisms, useful for diagnosing and treating lipid disorders,
PT   cardiovascular diseases and inflammatory diseases  -
XX
PS   Disclosure; Page 26-28; 41pp; English.
XX
CC   The present sequence is that of cDNA encoding the human adenosine
CC   triphosphate (ATP) binding cassette transporter 1 (ABC1) protein
CC   (see AAM50227).  The sequence includes an extended open reading
CC   frame (ORF) to that provided by the sequence in AAI70314, using
CC   an alternative ATG codon as initiation codon and thereby adding an
CC   extra 40 N-terminal amino acids to the encoded ABC1 protein (see
CC   AAM50228).  The invention provides 4 common polymorphisms in the
CC   ABC1 gene.  These were identified by sequencing the ABC1 gene in
CC   different Tangier kindreds.  In the variant genes (numbering as in
CC   AAI70314), G is changed to A at position 596, T is changed to C at
CC   position 1136, A is changed to G at position 2589 or G is changed
CC   to C at position 3456, or any combination of these.  All of these
CC   polymorphisms alter the amino acid sequence of ABC1 and therefore
CC   may affect its function.  The 2 most common polymorphisms (G596A)
CC   and A2589G) are both associated with a decreased in vitro ApoA-I
CC   mediated efflux of cholesterol from mononuclear phagocytes, a
CC   feature typical of Tangier disease.  3 Of the variants (G596A,
CC   A2589G and G3456C) are significantly increased in a population of
CC   men having low high density lipoprotein-cholesterol levels and
CC   established coronary heart disease (CHD) relative to CHD-free
CC   control subjects.  The use of the provided ABC1 polymorphisms for
CC   the diagnosis and treatment of lipid disorders, cardiovascular
CC   diseases, and inflammatory diseases (e.g. psoriasis, lupus
CC   erythematodes) is claimed.  Modulation of ABC1 transcripts or
CC   proteins by antisense or ribozyme technology or RNA decoys is also
CC   claimed.
XX
SQ   Sequence 7260 BP; 1834 A; 1765 C; 1905 G; 1756 T; 0 other;


  Query Match             100.0%;  Score 221;  DB 22;  Length 7260;
  Best Local Similarity   100.0%;  Pred. No. 3e-103;
  Matches  221;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 gtaattgcgagcgagagtgagtggggccgggacccgcagagccgagccgacccttctctc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8 gtaattgcgagcgagagtgagtggggccgggacccgcagagccgagccgacccttctctc 67

Qy     61 ccgggctgcggcagggcagggcggggagctccgcgcaccaacagagccggttctcagggc 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     68 ccgggctgcggcagggcagggcggggagctccgcgcaccaacagagccggttctcagggc 127

Qy    121 gctttgctccttgtttttttccccggttctgttttctccccttctccggaaggcttgtcaa 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    128 gctttgctccttgtttttttccccggttctgttttctccccttctccggaaggcttgtcaa 187
```

```
Qy    181 ggggtaggagaaagagacgcaaacacaaaagtggaaaacag 221
          |||||||||||||||||||||||||||||||||||||||||||
Db    188 ggggtaggagaaagagacgcaaacacaaaagtggaaaacag 228
```

RESULT 3
AAH04729
ID AAH04729 standard; cDNA; 763 BP.
XX
AC AAH04729;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:1564.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 1; SEQ ID 1564; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 763 BP; 137 A; 205 C; 260 G; 158 T; 3 other;

Query Match 99.1%; Score 219; DB 22; Length 763;
Best Local Similarity 100.0%; Pred. No. 3.2e-102;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      3 aattgcgagcgagagtgagtggggccgggacccgcagagccgagccgacccttctctccc 62
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 aattgcgagcgagagtgagtggggccgggacccgcagagccgagccgacccttctctccc 60

Qy     63 gggctgcggcagggcagggcggggagctccgcgcaccaacagagccggttctcagggcgc 122
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 gggctgcggcagggcagggcggggagctccgcgcaccaacagagccggttctcagggcgc 120

Qy    123 tttgctccttgtttttttccccggttctgttttctccccttctccggaaggcttgtcaagg 182
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 tttgctccttgtttttttccccggttctgttttctccccttctccggaaggcttgtcaagg 180

Qy    183 ggtaggagaaagagacgcaaacacaaaagtggaaaacag 221
          |||||||||||||||||||||||||||||||||||||||
Db    181 ggtaggagaaagagacgcaaacacaaaagtggaaaacag 219
```

RESULT 4
AAH17451
ID AAH17451 standard; cDNA; 1750 BP.
XX
AC AAH17451;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:16905.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 16905; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and

```
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 1750 BP; 291 A; 489 C; 586 G; 384 T; 0 other;


  Query Match             99.1%;  Score 219;  DB 22;  Length 1750;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-102;
  Matches  219;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 aattgcgagcgagagtgagtggggccgggacccgcagagccgagccgaccccttctctccc 62
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 aattgcgagcgagagtgagtggggccgggacccgcagagccgagccgaccccttctctccc 60

Qy      63 gggctgcggcagggcagggcggggagctccgcgcaccaacagagccggttctcagggcgc 122
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 gggctgcggcagggcagggcggggagctccgcgcaccaacagagccggttctcagggcgc 120

Qy     123 tttgctccttgtttttttccccggttctgttttctcccccttctccggaaggcttgtcaagg 182
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 tttgctccttgtttttttccccggttctgttttctcccccttctccggaaggcttgtcaagg 180

Qy     183 ggtaggagaaagagacgcaaacacaaaagtggaaaacag 221
          |||||||||||||||||||||||||||||||||||||||
Db     181 ggtaggagaaagagacgcaaacacaaaagtggaaaacag 219


RESULT   5
AAH07432
ID   AAH07432 standard; cDNA; 736 BP.
XX
AC   AAH07432;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA clone (5'-primer) SEQ ID NO:4267.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
PR   27-AUG-1999;   99JP-0300253.
PR   11-JAN-2000; 2000JP-0118776.
PR   02-MAY-2000; 2000JP-0183767.
PR   09-JUN-2000; 2000JP-0241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602
PT   full-length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full-length cDNAs -
XX
PS   Claim 1; SEQ ID 4267; 2537pp + CD ROM; English.
XX
CC   The present invention describes primer sets for synthesising 5602
CC   full-length cDNAs defined in the specification. Where a primer set
CC   comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC   to the complementary strand of a polynucleotide which comprises one of
CC   the 5602 nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 736 BP; 163 A; 199 C; 199 G; 170 T; 5 other;


  Query Match             98.2%;  Score 217;  DB 22;  Length 736;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-101;
  Matches  217;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 ttgcgagcgagagtgagtggggccgggacccgcagagccgagccgaccccttctctcccgg 64
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5 ttgcgagcgagagtgagtggggccgggacccgcagagccgagccgaccccttctctcccgg 64

Qy      65 gctgcggcagggcagggcggggagctccgcgcaccaacagagccggttctcagggcgctt 124
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65 gctgcggcagggcagggcggggagctccgcgcaccaacagagccggttctcagggcgctt 124

Qy     125 tgctccttgtttttttccccggttctgttttctcccccttctccggaaggcttgtcaagggg 184
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125 tgctccttgtttttttccccggttctgttttctcccccttctccggaaggcttgtcaagggg 184

Qy     185 taggagaaagagacgcaaacacaaaagtggaaaacag 221
          |||||||||||||||||||||||||||||||||||||
Db     185 taggagaaagagacgcaaacacaaaagtggaaaacag 221


RESULT   6
AAH18606
ID   AAH18606 standard; cDNA; 1556 BP.
XX
AC   AAH18606;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA sequence SEQ ID NO:18808.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
PR   27-AUG-1999;   99JP-0300253.
PR   11-JAN-2000; 2000JP-0118776.
PR   02-MAY-2000; 2000JP-0183767.
PR   09-JUN-2000; 2000JP-0241899.
```

XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 18808; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 1556 BP; 380 A; 363 C; 399 G; 414 T; 0 other;

```
  Query Match             98.2%;  Score 217;  DB 22;  Length 1556;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-101;
  Matches  217;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 ttgcgagcgagagtgagtggggccgggacccgcagagccgagccgacccttctctcccgg 64
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5 ttgcgagcgagagtgagtggggccgggacccgcagagccgagccgacccttctctcccgg 64

Qy      65 gctgcggcagggcagggcggggagctccgcgcaccaacagagccggttctcagggcgctt 124
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65 gctgcggcagggcagggcggggagctccgcgcaccaacagagccggttctcagggcgctt 124

Qy     125 tgctccttgtttttttccccggttctgttttctccccttctccggaaggcttgtcaagggg 184
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125 tgctccttgtttttttccccggttctgttttctccccttctccggaaggcttgtcaagggg 184

Qy     185 taggagaaagagacgcaaacacaaaagtggaaaacag 221
           |||||||||||||||||||||||||||||||||||||
Db     185 taggagaaagagacgcaaacacaaaagtggaaaacag 221
```

RESULT 7
ABA09200
ID ABA09200 standard; cDNA; 7086 BP.
XX
AC ABA09200;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human ABCA1 homologue-encoding cDNA, SEQ ID NO:976.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnerary; antiulcer; ss.
XX
OS Homo sapiens.
XX
PN WO200157188-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US03800.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-457740/49.
DR P-PSDB; ABB11956.
XX
PT Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -
XX
PS Claim 1; Page 833-835; 1963pp; English.
XX
CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g.,myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a cDNA encoding a

```
CC   novel human polypeptide of the invention.
XX
SQ   Sequence 7086 BP; 1773 A; 1739 C; 1859 G; 1715 T; 0 other;


  Query Match          92.8%;  Score 205;  DB 22;  Length 7086;
  Best Local Similarity  100.0%;  Pred. No. 4.5e-95;
  Matches  205;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      17 gtgagtggggccgggacccgcagagccgagccgaccctctctcccgggctgcggcaggg 76
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7 gtgagtggggccgggacccgcagagccgagccgaccctctctcccgggctgcggcaggg 66

Qy      77 cagggcggggagctccgcgcaccaacagagccggttctcagggcgctttgctccttgttt 136
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67 cagggcggggagctccgcgcaccaacagagccggttctcagggcgctttgctccttgttt 126

Qy     137 tttccccggttctgttttctcccctttctccggaaggcttgtcaaggggtaggagaaagag 196
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     127 tttccccggttctgttttctcccctttctccggaaggcttgtcaaggggtaggagaaagag 186

Qy     197 acgcaaacacaaaagtggaaaacag 221
           |||||||||||||||||||||||||
Db     187 acgcaaacacaaaagtggaaaacag 211


RESULT   8
AAK52667
ID   AAK52667 standard; cDNA; 7086 BP.
XX
AC   AAK52667;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 2196.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW   vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW   tissue growth factor; immunomodulatory; cancer; leukaemia;
KW   nervous system disorder; arthritis; inflammation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157190-A2.
XX
PD   09-AUG-2001.
XX
PF   05-FEB-2001; 2001WO-US04098.
XX
PR   03-FEB-2000; 2000US-0496914.
PR   27-APR-2000; 2000US-0560875.
PR   20-JUN-2000; 2000US-0598075.
PR   19-JUL-2000; 2000US-0620325.
PR   01-SEP-2000; 2000US-0654936.
PR   15-SEP-2000; 2000US-0663561.
PR   20-OCT-2000; 2000US-0693325.
PR   30-NOV-2000; 2000US-0728422.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT,  Asundi V,  Zhou P,  Xu C,  Cao Y,  Ma Y;
PI   Zhao QA,  Wang D,  Wang J,  Zhang J,  Ren F,  Chen R,  Wang ZW;
PI   Xue AJ,  Yang Y,  Wejhrman T,  Goodrich R;
XX
DR   WPI; 2001-476283/51.
DR   P-PSDB; AAM79534.
XX
PT   Nucleic acids encoding polypeptides with cytokine-like activities,
PT   useful in diagnosis and gene therapy -
XX
PS   Claim 1; Page 4558-4560; 6221pp; English.
XX
CC   The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC   encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC   cytokine, cell proliferation or cell differentiation or which may induce
CC   production of other cytokines in other cell populations. The
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation.
CC   Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC   (AAM80020) are omitted as the relevant pages from the sequence listing
CC   were missing at the time of publication.
XX
SQ   Sequence 7086 BP; 1773 A; 1739 C; 1859 G; 1715 T; 0 other;


  Query Match          92.8%;  Score 205;  DB 22;  Length 7086;
  Best Local Similarity  100.0%;  Pred. No. 4.5e-95;
  Matches  205;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      17 gtgagtggggccgggacccgcagagccgagccgaccctctctcccgggctgcggcaggg 76
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7 gtgagtggggccgggacccgcagagccgagccgaccctctctcccgggctgcggcaggg 66

Qy      77 cagggcggggagctccgcgcaccaacagagccggttctcagggcgctttgctccttgttt 136
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67 cagggcggggagctccgcgcaccaacagagccggttctcagggcgctttgctccttgttt 126

Qy     137 tttccccggttctgttttctcccctttctccggaaggcttgtcaaggggtaggagaaagag 196
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     127 tttccccggttctgttttctcccctttctccggaaggcttgtcaaggggtaggagaaagag 186

Qy     197 acgcaaacacaaaagtggaaaacag 221
           |||||||||||||||||||||||||
Db     187 acgcaaacacaaaagtggaaaacag 211


RESULT   9
AAK51683
ID   AAK51683 standard; cDNA; 7281 BP.
XX
AC   AAK51683;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 228.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW   vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW   tissue growth factor; immunomodulatory; cancer; leukaemia;
KW   nervous system disorder; arthritis; inflammation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157190-A2.
XX
PD   09-AUG-2001.
XX
PF   05-FEB-2001; 2001WO-US04098.
XX
PR   03-FEB-2000; 2000US-0496914.
PR   27-APR-2000; 2000US-0560875.
PR   20-JUN-2000; 2000US-0598075.
PR   19-JUL-2000; 2000US-0620325.
PR   01-SEP-2000; 2000US-0654936.
PR   15-SEP-2000; 2000US-0663561.
PR   20-OCT-2000; 2000US-0693325.
PR   30-NOV-2000; 2000US-0728422.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT,  Asundi V,  Zhou P,  Xu C,  Cao Y,  Ma Y;
```

```
PI   Zhao QA,  Wang D,  Wang J,  Zhang J,  Ren F,  Chen R,  Wang ZW;
PI   Xue AJ,  Yang Y,  Wejhrman T,  Goodrich R;
XX
DR   WPI; 2001-476283/51.
DR   P-PSDB; AAM78550.
XX
PT   Nucleic acids encoding polypeptides with cytokine-like activities,
PT   useful in diagnosis and gene therapy -
XX
PS   Claim 1; Page 1086-1096; 6221pp; English.
XX
CC   The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC   encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC   cytokine, cell proliferation or cell differentiation or which may induce
CC   production of other cytokines in other cell populations. The
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation.
CC   Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC   (AAM80020) are omitted as the relevant pages from the sequence listing
CC   were missing at the time of publication.
XX
SQ   Sequence 7281 BP; 1831 A; 1773 C; 1915 G; 1762 T; 0 other;


  Query Match             92.8%;  Score 205;  DB 22;  Length 7281;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-95;
  Matches  205;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      17 gtgagtggggccgggacccgcagagccgagccgacccttctctcccgggctgcggcaggg 76
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      45 gtgagtggggccgggacccgcagagccgagccgacccttctctcccgggctgcggcaggg 104

Qy      77 cagggcggggagctccgcgcaccaacagagccggttctcagggcgctttgctccttgttt 136
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105 cagggcggggagctccgcgcaccaacagagccggttctcagggcgctttgctccttgttt 164

Qy     137 tttccccggttctgttttctcccctttctccggaaggcttgtcaaggggtaggagaaagag 196
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     165 tttccccggttctgttttctcccctttctccggaaggcttgtcaaggggtaggagaaagag 224

Qy     197 acgcaaacacaaaagtggaaaacag 221
           |||||||||||||||||||||||||
Db     225 acgcaaacacaaaagtggaaaacag 249


RESULT  10
AAS06121
ID   AAS06121 standard; cDNA; 9854 BP.
XX
AC   AAS06121;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human ABC1 DNA sequence #2.
XX
KW   Human; ABC1 gene; atherosclerosis; reverse transport; cholesterol;
KW   cardiovascular; neurological; Tangier disease;, LCAT deficiency;
KW   lecithin-cholesterol acetyltransferase; malaria; diabetes; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             298..7078
FT                   /*tag=  a
FT                   /product= "Human ABC1 protein"
XX
PN   WO200130848-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-EP10886.
XX
PR   26-OCT-1999;   99EP-0402668.
PR   01-MAR-2000; 2000US-0186260.
XX
PA   (AVET ) AVENTIS PHARMA SA.
XX
PI   Denefle P,  Rosier-Montus M,  Arnould-Reguigne I,  Prades C,  Naudin L;
PI   Lemoine C,  Duverger N,  Jaye M,  Searfoss GH,  Remaley A,  Brewer HB;
PI   Dean M;
XX
DR   WPI; 2001-316327/33.
DR   P-PSDB; AAU02176.
XX
PT   New human ABC1 nucleic acids and polypeptides for treating
PT   atherosclerosis, malaria and diabetes -
XX
PS   Claim 1; Page 209-213; 368pp; English.
XX
CC   The sequence represents the coding sequence #2 of human ABC1. The
CC   nucleic acid sequence, primers and probes derived from the ABC1 sequence,
CC   and polypeptides and vectors are useful for the prevention of
CC   atherosclerosis, in a subject affected by a dysfunction in the reverse
CC   transport of cholesterol. The polypeptide encoded by the ABC1 gene is
CC   useful for screening for an active ingredient for the prevention or
CC   treatment of a disease resulting from dysfunction in the reverse
CC   transport of cholesterol. The nucleic acids and polypeptides are also
CC   useful for treating and preventing cardiovascular and neurological
CC   pathologies, and other diseases e.g. Tangier disease, lecithin-
CC   cholesterol (LCAT) deficiency, malaria and diabetes.
XX
SQ   Sequence 9854 BP; 2665 A; 2219 C; 2334 G; 2635 T; 1 other;


  Query Match             92.8%;  Score 205;  DB 22;  Length 9854;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-95;
  Matches  205;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      17 gtgagtggggccgggacccgcagagccgagccgacccttctctcccgggctgcggcaggg 76
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 gtgagtggggccgggacccgcagagccgagccgacccttctctcccgggctgcggcaggg 60

Qy      77 cagggcggggagctccgcgcaccaacagagccggttctcagggcgctttgctccttgttt 136
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 cagggcggggagctccgcgcaccaacagagccggttctcagggcgctttgctccttgttt 120

Qy     137 tttccccggttctgttttctcccctttctccggaaggcttgtcaaggggtaggagaaagag 196
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 tttccccggttctgttttctcccctttctccggaaggcttgtcaaggggtaggagaaagag 180

Qy     197 acgcaaacacaaaagtggaaaacag 221
           |||||||||||||||||||||||||
Db     181 acgcaaacacaaaagtggaaaacag 205


RESULT  11
AAF92831
ID   AAF92831 standard; DNA; 183999 BP.
XX
AC   AAF92831;
XX
DT   17-MAY-2001  (first entry)
XX
DE   Human ABC1 genomic DNA.
XX
KW   High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ds.
XX
OS   Homo sapiens.
XX
PN   WO200115676-A2.
XX
```

PD 08-MAR-2001.
XX
PF 01-SEP-2000; 2000WO-IB01492.
XX
PR 01-SEP-1999; 99US-0151977.
PR 15-MAR-2000; 2000US-0526193.
PR 23-JUN-2000; 2000US-0213958.
XX
PA (UYBR-) UNIV BRITISH COLUMBIA.
PA (XENO-) XENON GENETICS INC.
XX
PI Hayden MR, Brooks-Wilson AR, Pimstone SN, Clee SM;
XX
DR WPI; 2001-244356/25.
XX
PT Treating a lower than normal high density lipoprotein-cholesterol
PT (HDL-C) level, a higher than normal triglyceride level, or a
PT cardiovascular disease, by administering a compound that modulates LXR-
PT or RXR-mediated transcriptional activity -
XX
PS Claim 8; Fig 1; 317pp; English.
XX
CC The present invention relates to a method for treating a patient
CC diagnosed as having a lower than normal high density
CC lipoprotein-cholesterol (HDL-C) level, a higher than normal
CC triglyceride level, or a cardiovascular disease, involving
CC administering a compound that modulates LXR- or RXR-mediated
CC transcriptional activity or ABC1 expression or activity.
CC The LXR gene product may be used in an assay to identify
CC compounds useful for the treatment of a disease or condition selected a
CC lower than normal HDL cholesterol level, a higher than normal
CC triglyceride level, and a cardiovascular disease.
XX
SQ Sequence 183999 BP; 49549 A; 37944 C; 41170 G; 54950 T; 386 other;

Query Match 91.0%; Score 201; DB 22; Length 183999;
Best Local Similarity 100.0%; Pred. No. 4.8e-93;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 gtaattgcgagcgagagtgagtggggccgggacccgcagagccgagccgacccttctctc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28790 gtaattgcgagcgagagtgagtggggccgggacccgcagagccgagccgacccttctctc 28849

Qy     61 ccgggctgcggcagggcagggcggggagctccgcgcaccaacagagccggttctcagggc 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28850 ccgggctgcggcagggcagggcggggagctccgcgcaccaacagagccggttctcagggc 28909

Qy    121 gctttgctccttgtttttttccccggttctgttttctccccttctccggaaggcttgtcaa 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28910 gctttgctccttgtttttttccccggttctgttttctccccttctccggaaggcttgtcaa 28969

Qy    181 ggggtaggagaaagagacgca 201
          |||||||||||||||||||||
Db  28970 ggggtaggagaaagagacgca 28990
```

RESULT 12
AAC09615
ID AAC09615 standard; cDNA; 227 BP.
XX
AC AAC09615;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 13690.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1; SEQ ID 13690; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 227 BP; 44 A; 65 C; 73 G; 45 T; 0 other;

Query Match 89.1%; Score 197; DB 21; Length 227;
Best Local Similarity 100.0%; Pred. No. 5.6e-91;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 gtaattgcgagcgagagtgagtggggccgggacccgcagagccgagccgacccttctctc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     31 gtaattgcgagcgagagtgagtggggccgggacccgcagagccgagccgacccttctctc 90

Qy     61 ccgggctgcggcagggcagggcggggagctccgcgcaccaacagagccggttctcagggc 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     91 ccgggctgcggcagggcagggcggggagctccgcgcaccaacagagccggttctcagggc 150

Qy    121 gctttgctccttgtttttttccccggttctgttttctccccttctccggaaggcttgtcaa 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    151 gctttgctccttgtttttttccccggttctgttttctccccttctccggaaggcttgtcaa 210

Qy    181 ggggtaggagaaagaga 197
          |||||||||||||||||
Db    211 ggggtaggagaaagaga 227
```

RESULT 13
AAF24680
ID AAF24680 standard; DNA; 10442 BP.
XX
AC AAF24680;
XX
DT 20-APR-2001 (first entry)
XX
DE Nucleotide sequence of a human ABC1 polypeptide.
XX
KW Human; adenosine triphosphate binding cassette protein 1; ABC1;
KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;
KW atherosclerosis; cholesterol transport; ss.
XX

```
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             291..7076
FT                   /*tag=  a
FT                   /product= "ABC1 polypeptide"
XX
PN   WO200078972-A2.
XX
PD   28-DEC-2000.
XX
PF   16-JUN-2000; 2000WO-US16765.
XX
PR   18-JUN-1999;   99US-0140264.
PR   14-SEP-1999;   99US-0153872.
PR   19-NOV-1999;   99US-0166573.
XX
PA   (CVTH-) CV THERAPEUTICS INC.
XX
PI   Lawn RM,  Wade D,  Garvin M;
XX
DR   WPI; 2001-137812/14.
XX
PT   Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,
PT   useful for the development of agents for the treatment of heart disease
PT   and other disorders associated with hypercholesterolemia and
PT   atherosclerosis -
XX
PS   Disclosure; Page 122-128; 215pp; English.
XX
CC   The present sequence encodes a human adenosine triphosphate (ATP)
CC   binding cassette protein (ABC) 1 polypeptide. ABC1 resides in cell
CC   membranes and utilises ATP hydrolysis to transport a wide variety of
CC   substrates across the plasma membrane. ABC1 is a pivotal protein in
CC   the apolipoprotein-mediated mobilisation of intracellular cholesterol
CC   stores. ABC1 is defective in Tangier disease, a genetic disorder
CC   characterised by abnormal HDL-cholesterol metabolism. The ABC1 gene is
CC   localised to chromosome 9q22-9q31. The ABC1 genes and proteins are
CC   useful for developing pharmaceutical agents for the treatment of heart
CC   disease and other disorders associated with hypercholesterolemia and
CC   atherosclerosis. The genes are useful for developing screening assays to
CC   screen for compounds that regulate the expression of genes associated
CC   with cholesterol transport. The genes and proteins are also useful for
CC   are also useful as diagnostic indicators of cardiovascular disease and
CC   other disorders associated with hypercholesterolemia.
XX
SQ   Sequence 10442 BP; 2898 A; 2297 C; 2408 G; 2835 T; 4 other;


  Query Match             89.1%;  Score 197;  DB 22;  Length 10442;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-91;
  Matches  197;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      25 ggccgggacccgcagagccgagccgaccccttctctcccgggctgcggcagggcagggcgg 84
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ggccgggacccgcagagccgagccgaccccttctctcccgggctgcggcagggcagggcgg 60

Qy      85 ggagctccgcgcaccaacagagccggttctcagggcgctttgctccttgtttttttccccg 144
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ggagctccgcgcaccaacagagccggttctcagggcgctttgctccttgtttttttccccg 120

Qy     145 gttctgttttctccccttctccggaaggcttgtcaaggggtaggagaaagagacgcaaac 204
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 gttctgttttctccccttctccggaaggcttgtcaaggggtaggagaaagagacgcaaac 180

Qy     205 acaaaagtggaaaacag 221
           |||||||||||||||||
Db     181 acaaaagtggaaaacag 197


RESULT  14
AAF24702
ID   AAF24702 standard; DNA; 10442 BP.
XX
AC   AAF24702;
XX
DT   20-APR-2001  (first entry)
XX
DE   Nucleotide sequence of a human ABC1 polypeptide.
XX
KW   Human; adenosine triphosphate binding cassette protein 1; ABC1;
KW   apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW   chromosome 9q22-9q31; heart disease; hypercholesterolemia;
KW   atherosclerosis; cholesterol transport; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             291..7076
FT                   /*tag=  a
FT                   /product= "ABC1 polypeptide"
XX
PN   WO200078971-A2.
XX
PD   28-DEC-2000.
XX
PF   16-JUN-2000; 2000WO-US16591.
XX
PR   18-JUN-1999;   99US-0140264.
PR   14-SEP-1999;   99US-0153872.
PR   19-NOV-1999;   99US-0166573.
XX
PA   (CVTH-) CV THERAPEUTICS INC.
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Lawn RM,  Wade D,  Oram JF,  Garvin M;
XX
DR   WPI; 2001-137811/14.
DR   P-PSDB; AAB31365.
XX
PT   Adenosine triphosphate (ATP) binding cassette protein (ABC) 1
PT   polynucleotides and polypeptides, useful for treatment of heart disease
PT   and other disorders associated with hypercholesterolemia and
PT   atherosclerosis -
XX
PS   Claim 3; Page 117-123; 211pp; English.
XX
CC   The present sequence encodes a human adenosine triphosphate (ATP)
CC   binding cassette protein (ABC) 1 polypeptide. ABC1 resides in cell
CC   membranes and utilises ATP hydrolysis to transport a wide variety of
CC   substrates across the plasma membrane. ABC1 is a pivotal protein in
CC   the apolipoprotein-mediated mobilisation of intracellular cholesterol
CC   stores. ABC1 is defective in Tangier disease, a genetic disorder
CC   characterised by abnormal HDL-cholesterol metabolism. The ABC1 gene is
CC   localised to chromosome 9q22-9q31. The ABC1 genes and proteins are
CC   useful for developing pharmaceutical agents for the treatment of heart
CC   disease and other disorders associated with hypercholesterolemia and
CC   atherosclerosis. The genes are useful for developing screening assays to
CC   screen for compounds that regulate the expression of genes associated
CC   with cholesterol transport. The genes and proteins are also useful for
CC   are also useful as diagnostic indicators of cardiovascular disease and
CC   other disorders associated with hypercholesterolemia.
XX
SQ   Sequence 10442 BP; 2898 A; 2297 C; 2408 G; 2835 T; 4 other;


  Query Match             89.1%;  Score 197;  DB 22;  Length 10442;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-91;
  Matches  197;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      25 ggccgggacccgcagagccgagccgaccccttctctcccgggctgcggcagggcagggcgg 84
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ggccgggacccgcagagccgagccgaccccttctctcccgggctgcggcagggcagggcgg 60

Qy      85 ggagctccgcgcaccaacagagccggttctcagggcgctttgctccttgtttttttccccg 144
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     61 ggagctccgcgcaccaacagagccggttctcagggcgctttgctccttgtttttttccccg 120

Qy    145 gttctgttttctcccccttctccggaaggcttgtcaaggggtaggagaaagagacgcaaac 204
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 gttctgttttctcccccttctccggaaggcttgtcaaggggtaggagaaagagacgcaaac 180

Qy    205 acaaaagtggaaaacag 221
          |||||||||||||||||
Db    181 acaaaagtggaaaacag 197


RESULT  15
AAF24685
ID   AAF24685 standard; DNA; 10474 BP.
XX
AC   AAF24685;
XX
DT   20-APR-2001  (first entry)
XX
DE   Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.
XX
KW   Human; adenosine triphosphate binding cassette protein 1; ABC1;
KW   apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW   chromosome 9q22-9q31; heart disease; hypercholesterolemia;
KW   atherosclerosis; cholesterol transport; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             323..7108
FT                   /*tag=  a
FT                   /product= "defective ABC1 polypeptide"
XX
PN   WO200078972-A2.
XX
PD   28-DEC-2000.
XX
PF   16-JUN-2000; 2000WO-US16765.
XX
PR   18-JUN-1999;   99US-0140264.
PR   14-SEP-1999;   99US-0153872.
PR   19-NOV-1999;   99US-0166573.
XX
PA   (CVTH-) CV THERAPEUTICS INC.
XX
PI   Lawn RM,  Wade D,  Garvin M;
XX
DR   WPI; 2001-137812/14.
XX
PT   Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,
PT   useful for the development of agents for the treatment of heart disease
PT   and other disorders associated with hypercholesterolemia and
PT   atherosclerosis -
XX
PS   Disclosure; Page 148-154; 215pp; English.
XX
CC   The present sequence encodes a human adenosine triphosphate (ATP)
CC   binding cassette protein (ABC) 1 polypeptide, and is isolated from
CC   a Tangier disease patient. ABC1 resides in cell membranes and utilises
CC   ATP hydrolysis to transport a wide variety of substrates across the
CC   plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated
CC   mobilisation of intracellular cholesterol stores. ABC1 is defective in
CC   Tangier disease, a genetic disorder characterised by abnormal
CC   HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome
CC   9q22-9q31. The ABC1 genes and proteins are useful for developing
CC   pharmaceutical agents for the treatment of heart disease and other
CC   disorders associated with hypercholesterolemia and atherosclerosis. The
CC   genes are useful for developing screening assays to screen for compounds
CC   that regulate the expression of genes associated with cholesterol
CC   transport. The genes and proteins are also useful for are also useful
CC   as diagnostic indicators of cardiovascular disease and other disorders
CC   associated with hypercholesterolemia.
XX

SQ   Sequence 10474 BP; 2906 A; 2305 C; 2416 G; 2843 T; 4 other;


  Query Match             85.1%;  Score 188;  DB 22;  Length 10474;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-86;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     34 ccgcagagccgagccgacccttctctcccgggctgcggcagggcagggcggggagctccg 93
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     42 ccgcagagccgagccgacccttctctcccgggctgcggcagggcagggcggggagctccg 101

Qy     94 cgcaccaacagagccggttctcagggcgctttgctccttgtttttttccccggttctgttt 153
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    102 cgcaccaacagagccggttctcagggcgctttgctccttgtttttttccccggttctgttt 161

Qy    154 tctcccccttctccggaaggcttgtcaaggggtaggagaaagagacgcaaacacaaaagtg 213
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    162 tctcccccttctccggaaggcttgtcaaggggtaggagaaagagacgcaaacacaaaagtg 221

Qy    214 gaaaacag 221
          ||||||||
Db    222 gaaaacag 229


RESULT  16
AAF24686
ID   AAF24686 standard; DNA; 10474 BP.
XX
AC   AAF24686;
XX
DT   20-APR-2001  (first entry)
XX
DE   Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.
XX
KW   Human; adenosine triphosphate binding cassette protein 1; ABC1;
KW   apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW   chromosome 9q22-9q31; heart disease; hypercholesterolemia;
KW   atherosclerosis; cholesterol transport; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             323..7108
FT                   /*tag=  a
FT                   /product= "defective ABC1 polypeptide"
XX
PN   WO200078972-A2.
XX
PD   28-DEC-2000.
XX
PF   16-JUN-2000; 2000WO-US16765.
XX
PR   18-JUN-1999;   99US-0140264.
PR   14-SEP-1999;   99US-0153872.
PR   19-NOV-1999;   99US-0166573.
XX
PA   (CVTH-) CV THERAPEUTICS INC.
XX
PI   Lawn RM,  Wade D,  Garvin M;
XX
DR   WPI; 2001-137812/14.
XX
PT   Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,
PT   useful for the development of agents for the treatment of heart disease
PT   and other disorders associated with hypercholesterolemia and
PT   atherosclerosis -
XX
PS   Disclosure; Page 170-176; 215pp; English.
XX
CC   The present sequence encodes a human adenosine triphosphate (ATP)
CC   binding cassette protein (ABC) 1 polypeptide, and is isolated from
CC   a Tangier disease patient. ABC1 resides in cell membranes and utilises
CC   ATP hydrolysis to transport a wide variety of substrates across the
```

```
CC   plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated
CC   mobilisation of intracellular cholesterol stores. ABC1 is defective in
CC   Tangier disease, a genetic disorder characterised by abnormal
CC   HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome
CC   9q22-9q31. The ABC1 genes and proteins are useful for developing
CC   pharmaceutical agents for the treatment of heart disease and other
CC   disorders associated with hypercholesterolemia and atherosclerosis. The
CC   genes are useful for developing screening assays to screen for compounds
CC   that regulate the expression of genes associated with cholesterol
CC   transport. The genes and proteins are also useful for are also useful
CC   as diagnostic indicators of cardiovascular disease and other disorders
CC   associated with hypercholesterolemia.
XX
SQ   Sequence 10474 BP; 2907 A; 2304 C; 2415 G; 2844 T; 4 other;


  Query Match             85.1%;  Score 188;  DB 22;  Length 10474;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-86;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      34 ccgcagagccgagccgacccttctctcccgggctgcggcagggcagggcggggagctccg 93
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42 ccgcagagccgagccgacccttctctcccgggctgcggcagggcagggcggggagctccg 101

Qy      94 cgcaccaacagagccggttctcagggcgctttgctccttgtttttttccccggttctgttt 153
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102 cgcaccaacagagccggttctcagggcgctttgctccttgtttttttccccggttctgttt 161

Qy     154 tctccccttctccggaaggcttgtcaaggggtaggagaaagagacgcaaacacaaaagtg 213
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     162 tctccccttctccggaaggcttgtcaaggggtaggagaaagagacgcaaacacaaaagtg 221

Qy     214 gaaaacag 221
           ||||||||
Db     222 gaaaacag 229


RESULT  17
AAF24707
ID   AAF24707 standard; DNA; 10474 BP.
XX
AC   AAF24707;
XX
DT   20-APR-2001  (first entry)
XX
DE   Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.
XX
KW   Human; adenosine triphosphate binding cassette protein 1; ABC1;
KW   apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW   chromosome 9q22-9q31; heart disease; hypercholesterolemia;
KW   atherosclerosis; cholesterol transport; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             323..7108
FT                   /*tag=  a
FT                   /product= "defective ABC1 polypeptide"
XX
PN   WO200078971-A2.
XX
PD   28-DEC-2000.
XX
PF   16-JUN-2000; 2000WO-US16591.
XX
PR   18-JUN-1999;   99US-0140264.
PR   14-SEP-1999;   99US-0153872.
PR   19-NOV-1999;   99US-0166573.
XX
PA   (CVTH-) CV THERAPEUTICS INC.
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Lawn RM,  Wade D,  Oram JF,  Garvin M;
XX
DR   WPI; 2001-137811/14.
DR   P-PSDB; AAB31366.
XX
PT   Adenosine triphosphate (ATP) binding cassette protein (ABC) 1
PT   polynucleotides and polypeptides, useful for treatment of heart disease
PT   and other disorders associated with hypercholesterolemia and
PT   atherosclerosis -
XX
PS   Claim 27; Page 144-150; 211pp; English.
XX
CC   The present sequence encodes a human adenosine triphosphate (ATP)
CC   binding cassette protein (ABC) 1 polypeptide, and is isolated from
CC   a Tangier disease patient. ABC1 resides in cell membranes and utilises
CC   ATP hydrolysis to transport a wide variety of substrates across the
CC   plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated
CC   mobilisation of intracellular cholesterol stores. ABC1 is defective in
CC   Tangier disease, a genetic disorder characterised by abnormal
CC   HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome
CC   9q22-9q31. The ABC1 genes and proteins are useful for developing
CC   pharmaceutical agents for the treatment of heart disease and other
CC   disorders associated with hypercholesterolemia and atherosclerosis. The
CC   genes are useful for developing screening assays to screen for compounds
CC   that regulate the expression of genes associated with cholesterol
CC   transport. The genes and proteins are also useful for are also useful
CC   as diagnostic indicators of cardiovascular disease and other disorders
CC   associated with hypercholesterolemia.
XX
SQ   Sequence 10474 BP; 2906 A; 2305 C; 2416 G; 2843 T; 4 other;


  Query Match             85.1%;  Score 188;  DB 22;  Length 10474;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-86;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      34 ccgcagagccgagccgacccttctctcccgggctgcggcagggcagggcggggagctccg 93
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42 ccgcagagccgagccgacccttctctcccgggctgcggcagggcagggcggggagctccg 101

Qy      94 cgcaccaacagagccggttctcagggcgctttgctccttgtttttttccccggttctgttt 153
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102 cgcaccaacagagccggttctcagggcgctttgctccttgtttttttccccggttctgttt 161

Qy     154 tctccccttctccggaaggcttgtcaaggggtaggagaaagagacgcaaacacaaaagtg 213
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     162 tctccccttctccggaaggcttgtcaaggggtaggagaaagagacgcaaacacaaaagtg 221

Qy     214 gaaaacag 221
           ||||||||
Db     222 gaaaacag 229


RESULT  18
AAF24708
ID   AAF24708 standard; DNA; 10474 BP.
XX
AC   AAF24708;
XX
DT   20-APR-2001  (first entry)
XX
DE   Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.
XX
KW   Human; adenosine triphosphate binding cassette protein 1; ABC1;
KW   apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW   chromosome 9q22-9q31; heart disease; hypercholesterolemia;
KW   atherosclerosis; cholesterol transport; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             323..7108
FT                   /*tag=  a
FT                   /product= "defective ABC1 polypeptide"
```

```
XX
PN   WO200078971-A2.
XX
PD   28-DEC-2000.
XX
PF   16-JUN-2000; 2000WO-US16591.
XX
PR   18-JUN-1999;   99US-0140264.
PR   14-SEP-1999;   99US-0153872.
PR   19-NOV-1999;   99US-0166573.
XX
PA   (CVTH-) CV THERAPEUTICS INC.
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Lawn RM,  Wade D,  Oram JF,  Garvin M;
XX
DR   WPI; 2001-137811/14.
DR   P-PSDB; AAB31367.
XX
PT   Adenosine triphosphate (ATP) binding cassette protein (ABC) 1
PT   polynucleotides and polypeptides, useful for treatment of heart disease
PT   and other disorders associated with hypercholesterolemia and
PT   atherosclerosis -
XX
PS   Claim 30; Page 165-172; 211pp; English.
XX
CC   The present sequence encodes a human adenosine triphosphate (ATP)
CC   binding cassette protein (ABC) 1 polypeptide, and is isolated from
CC   a Tangier disease patient. ABC1 resides in cell membranes and utilises
CC   ATP hydrolysis to transport a wide variety of substrates across the
CC   plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated
CC   mobilisation of intracellular cholesterol stores. ABC1 is defective in
CC   Tangier disease, a genetic disorder characterised by abnormal
CC   HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome
CC   9q22-9q31. The ABC1 genes and proteins are useful for developing
CC   pharmaceutical agents for the treatment of heart disease and other
CC   disorders associated with hypercholesterolemia and atherosclerosis. The
CC   genes are useful for developing screening assays to screen for compounds
CC   that regulate the expression of genes associated with cholesterol
CC   transport. The genes and proteins are also useful for are also useful
CC   as diagnostic indicators of cardiovascular disease and other disorders
CC   associated with hypercholesterolemia.
XX
SQ   Sequence 10474 BP; 2907 A; 2304 C; 2415 G; 2844 T; 4 other;


  Query Match             85.1%;  Score 188;  DB 22;  Length 10474;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-86;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      34 ccgcagagccgagccgacccttctctcccgggctgcggcagggcagggcggggagctccg 93
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42 ccgcagagccgagccgacccttctctcccgggctgcggcagggcagggcggggagctccg 101

Qy      94 cgcaccaacagagccggttctcagggcgctttgctccttgttttttccccggttctgttt 153
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102 cgcaccaacagagccggttctcagggcgctttgctccttgttttttccccggttctgttt 161

Qy     154 tctccccttctccggaaggcttgtcaaggggtaggagaaagagacgcaaacacaaaagtg 213
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     162 tctccccttctccggaaggcttgtcaaggggtaggagaaagagacgcaaacacaaaagtg 221

Qy     214 gaaaacag 221
           ||||||||
Db     222 gaaaacag 229


RESULT  19
AAS04035
ID   AAS04035 standard; cDNA; 446 BP.
XX
AC   AAS04035;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Partial human ABC1 cDNA sequence.
XX
KW   Human; ABC1 gene; atherosclerosis; reverse transport; cholesterol;
KW   cardiovascular; neurological; Tangier disease;, LCAT deficiency;
KW   lecithin-cholesterol acetyltransferase; malaria; diabetes; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             185..438
FT                   /*tag=  a
FT                   /product= "Human ABC1 protein, amino acids 1 to 60"
XX
PN   WO200130848-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-EP10886.
XX
PR   26-OCT-1999;   99EP-0402668.
PR   01-MAR-2000; 2000US-0186260.
XX
PA   (AVET ) AVENTIS PHARMA SA.
XX
PI   Denefle P,  Rosier-Montus M,  Arnould-Reguigne I,  Prades C,  Naudin L;
PI   Lemoine C,  Duverger N,  Jaye M,  Searfoss GH,  Remaley A,  Brewer HB;
PI   Dean M;
XX
DR   WPI; 2001-316327/33.
DR   P-PSDB; AAU02176.
XX
PT   New human ABC1 nucleic acids and polypeptides for treating
PT   atherosclerosis, malaria and diabetes -
XX
PS   Example 2; Page 167; 368pp; English.
XX
CC   The sequence represents the partial coding sequence of human ABC1,
CC   which encodes amino acids 1-60 of the human ABC1 protein. The nucleic
CC   acid sequence, primers and probes derived from the ABC1 sequence, and
CC   polypeptides and vectors are useful for the prevention of
CC   atherosclerosis, in a subject affected by a dysfunction in the reverse
CC   transport of cholesterol. The polypeptide encoded by the ABC1 gene is
CC   useful for screening for an active ingredient for the prevention or
CC   treatment of a disease resulting from dysfunction in the reverse
CC   transport of cholesterol. The nucleic acids and polypeptides are also
CC   useful for treating and preventing cardiovascular and neurological
CC   pathologies, and other diseases e.g. Tangier disease, lecithin-
CC   cholesterol (LCAT) deficiency, malaria and diabetes.
XX
SQ   Sequence 446 BP; 96 A; 123 C; 112 G; 115 T; 0 other;


  Query Match             41.6%;  Score 92;  DB 22;  Length 446;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-37;
  Matches   92;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     130 cttgtttttttccccggttctgttttctccccttctccggaaggcttgtcaaggggtagga 189
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 cttgtttttttccccggttctgttttctccccttctccggaaggcttgtcaaggggtagga 60

Qy     190 gaaagagacgcaaacacaaaagtggaaaacag 221
           ||||||||||||||||||||||||||||||||
Db      61 gaaagagacgcaaacacaaaagtggaaaacag 92


RESULT  20
AAS06120
ID   AAS06120 standard; cDNA; 9741 BP.
XX
AC   AAS06120;
XX
```

```
DT   12-SEP-2001  (first entry)
XX
DE   Human ABC1 DNA sequence #1.
XX
KW   Human; ABC1 gene; atherosclerosis; reverse transport; cholesterol;
KW   cardiovascular; neurological; Tangier disease;, LCAT deficiency;
KW   lecithin-cholesterol acetyltransferase; malaria; diabetes; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             185..6967
FT                   /*tag=  a
FT                   /product= "Human ABC1 protein"
XX
PN   WO200130848-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-EP10886.
XX
PR   26-OCT-1999;   99EP-0402668.
PR   01-MAR-2000; 2000US-0186260.
XX
PA   (AVET ) AVENTIS PHARMA SA.
XX
PI   Denefle P,  Rosier-Montus M,  Arnould-Reguigne I,  Prades C,  Naudin L;
PI   Lemoine C,  Duverger N,  Jaye M,  Searfoss GH,  Remaley A,  Brewer HB;
PI   Dean M;
XX
DR   WPI; 2001-316327/33.
DR   P-PSDB; AAU02176.
XX
PT   New human ABC1 nucleic acids and polypeptides for treating
PT   atherosclerosis, malaria and diabetes -
XX
PS   Claim 1; Page 204-208; 368pp; English.
XX
CC   The sequence represents the coding sequence #1 of human ABC1. The
CC   nucleic acid sequence, primers and probes derived from the ABC1 sequence,
CC   and polypeptides and vectors are useful for the prevention of
CC   atherosclerosis, in a subject affected by a dysfunction in the reverse
CC   transport of cholesterol. The polypeptide encoded by the ABC1 gene is
CC   useful for screening for an active ingredient for the prevention or
CC   treatment of a disease resulting from dysfunction in the reverse
CC   transport of cholesterol. The nucleic acids and polypeptides are also
CC   useful for treating and preventing cardiovascular and neurological
CC   pathologies, and other diseases e.g. Tangier disease, lecithin-
CC   cholesterol (LCAT) deficiency, malaria and diabetes.
XX
SQ   Sequence 9741 BP; 2650 A; 2180 C; 2290 G; 2620 T; 1 other;

  Query Match             41.6%;  Score 92;  DB 22;  Length 9741;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-37;
  Matches   92;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    130 cttgtttttttccccggttctgttttctccccttctccggaaggcttgtcaaggggtagga 189
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 cttgtttttttccccggttctgttttctccccttctccggaaggcttgtcaaggggtagga 60

Qy    190 gaaagagacgcaaacacaaaagtggaaaacag 221
          |||||||||||||||||||||||||||||||
Db     61 gaaagagacgcaaacacaaaagtggaaaacag 92


RESULT  21
AAF24681
ID   AAF24681 standard; DNA; 1643 BP.
XX
AC   AAF24681;
XX
DT   20-APR-2001  (first entry)
XX
DE   Nucleotide sequence of the 5' flanking region of the human ABC1 gene.
XX
KW   Human; adenosine triphosphate binding cassette protein 1; ABC1;
KW   apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW   chromosome 9q22-9q31; heart disease; hypercholesterolemia;
KW   atherosclerosis; cholesterol transport; ss.
XX
OS   Homo sapiens.
XX
PN   WO200078972-A2.
XX
PD   28-DEC-2000.
XX
PF   16-JUN-2000; 2000WO-US16765.
XX
PR   18-JUN-1999;   99US-0140264.
PR   14-SEP-1999;   99US-0153872.
PR   19-NOV-1999;   99US-0166573.
XX
PA   (CVTH-) CV THERAPEUTICS INC.
XX
PI   Lawn RM,  Wade D,  Garvin M;
XX
DR   WPI; 2001-137812/14.
XX
PT   Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,
PT   useful for the development of agents for the treatment of heart disease
PT   and other disorders associated with hypercholesterolemia and
PT   atherosclerosis -
XX
PS   Claim 1; Page 143-144; 215pp; English.
XX
CC   The present sequence represents the 5' flanking region of the huuman
CC   adenosine triphosphate (ATP) binding cassette protein (ABC) 1 gene. ABC1
CC   resides in cell membranes and utilises ATP hydrolysis to transport a wide
CC   variety of substrates across the plasma membrane. ABC1 is a pivotal
CC   protein in the apolipoprotein-mediated mobilisation of intracellular
CC   cholesterol stores. ABC1 is defective in Tangier disease, a genetic
CC   disorder characterised by abnormal HDL-cholesterol metabolism. The ABC1
CC   gene is localised to chromosome 9q22-9q31. The ABC1 genes and proteins
CC   are useful for developing pharmaceutical agents for the treatment of
CC   heart disease and other disorders associated with hypercholesterolemia
CC   and atherosclerosis. The genes are useful for developing screening assays
CC   to screen for compounds that regulate the expression of genes associated
CC   with cholesterol transport. The genes and proteins are also useful for
CC   are also useful as diagnostic indicators of cardiovascular disease and
CC   other disorders associated with hypercholesterolemia.
XX
SQ   Sequence 1643 BP; 370 A; 413 C; 457 G; 403 T; 0 other;

  Query Match             41.2%;  Score 91;  DB 22;  Length 1643;
  Best Local Similarity   100.0%;  Pred. No. 8e-37;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 gtaattgcgagcgagagtgagtggggccgggacccgcagagccgagccgacccttctctc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1553 gtaattgcgagcgagagtgagtggggccgggacccgcagagccgagccgacccttctctc 1612

Qy     61 ccgggctgcggcagggcagggcggggagctc 91
          |||||||||||||||||||||||||||||||
Db   1613 ccgggctgcggcagggcagggcggggagctc 1643


RESULT  22
AAF24703
ID   AAF24703 standard; DNA; 1643 BP.
XX
AC   AAF24703;
XX
DT   20-APR-2001  (first entry)
XX
```

DE Nucleotide sequence of the 5' flanking region of the human ABC1 gene.
XX
KW Human; adenosine triphosphate binding cassette protein 1; ABC1;
KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;
KW atherosclerosis; cholesterol transport; ss.
XX
OS Homo sapiens.
XX
PN WO200078971-A2.
XX
PD 28-DEC-2000.
XX
PF 16-JUN-2000; 2000WO-US16591.
XX
PR 18-JUN-1999; 99US-0140264.
PR 14-SEP-1999; 99US-0153872.
PR 19-NOV-1999; 99US-0166573.
XX
PA (CVTH-) CV THERAPEUTICS INC.
PA (UNIW) UNIV WASHINGTON.
XX
PI Lawn RM, Wade D, Oram JF, Garvin M;
XX
DR WPI; 2001-137811/14.
XX
PT Adenosine triphosphate (ATP) binding cassette protein (ABC) 1
PT polynucleotides and polypeptides, useful for treatment of heart disease
PT and other disorders associated with hypercholesterolemia and
PT atherosclerosis -
XX
PS Disclosure; Page 138-139; 211pp; English.
XX
CC The present sequence represents the 5' flanking region of the huuman
CC adenosine triphosphate (ATP) binding cassette protein (ABC) 1 gene. ABC1
CC resides in cell membranes and utilises ATP hydrolysis to transport a wide
CC variety of substrates across the plasma membrane. ABC1 is a pivotal
CC protein in the apolipoprotein-mediated mobilisation of intracellular
CC cholesterol stores. ABC1 is defective in Tangier disease, a genetic
CC disorder characterised by abnormal HDL-cholesterol metabolism. The ABC1
CC gene is localised to chromosome 9q22-9q31. The ABC1 genes and proteins
CC are useful for developing pharmaceutical agents for the treatment of
CC heart disease and other disorders associated with hypercholesterolemia
CC and atherosclerosis. The genes are useful for developing screening assays
CC to screen for compounds that regulate the expression of genes associated
CC with cholesterol transport. The genes and proteins are also useful for
CC are also useful as diagnostic indicators of cardiovascular disease and
CC other disorders associated with hypercholesterolemia.
XX
SQ Sequence 1643 BP; 370 A; 413 C; 457 G; 403 T; 0 other;

```
  Query Match             41.2%;  Score 91;  DB 22;  Length 1643;
  Best Local Similarity   100.0%;  Pred. No. 8e-37;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 gtaattgcgagcgagagtgagtggggccgggacccgcagagccgagccgacccttctctc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1553 gtaattgcgagcgagagtgagtggggccgggacccgcagagccgagccgacccttctctc 1612

Qy      61 ccgggctgcggcagggcagggcggggagctc 91
           |||||||||||||||||||||||||||||||
Db    1613 ccgggctgcggcagggcagggcggggagctc 1643
```

RESULT 23
AAC00305/c
ID AAC00305 standard; cDNA; 310 BP.
XX
AC AAC00305;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 303.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
DR P-PSDB; AAG00299.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1; SEQ ID 303; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.
XX
SQ Sequence 310 BP; 67 A; 82 C; 93 G; 68 T; 0 other;

```
  Query Match             8.6%;  Score 19;  DB 21;  Length 310;
  Best Local Similarity   100.0%;  Pred. No. 5;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     178 caaggggtaggagaaagag 196
           ||||||||||||||||||
Db     132 CAAGGGGTAGGAGAAAGAG 114
```

RESULT 24
AAV90187
ID AAV90187 standard; cDNA; 447 BP.
XX
AC AAV90187;
XX
DT 15-FEB-1999 (first entry)
XX
DE EST clone DH318.
XX
KW Human; secreted protein; expressed sequence tag; EST; haematopoiesis;
KW tissue growth; activin; inhibin; chemotaxis; chemokinesis; haemostatic;
KW receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;
KW gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO9845436-A2.

XX
PD 15-OCT-1998.
XX
PF 10-APR-1998; 98WO-US06955.
XX
PR 10-APR-1997; 97US-0838821.
XX
PA (GEMY) GENETICS INST INC.
XX
PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
PI Racie LA, Spaulding V, Treacy M;
XX
DR WPI; 1999-070077/06.
XX
PT New polynucleotides encoding human secreted proteins - derived from
PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
PT ovary, pituitary, retina and colon cDNA libraries.
XX
PS Claim 1; Page 459; 618pp; English.
XX
CC The present sequence represents a human expressed sequence tag (EST).
CC The polynucleotide, which is a secreted EST, and the encoded protein
CC are predicted to have useful biological activities which would make
CC them suitable for treating, preventing or ameliorating medical
CC conditions in humans and animals, although no supporting data is
CC given. Suggested activities include nutritional activity, immune
CC stimulating or suppressing activity, haematopoiesis regulating
CC activity, tissue growth activity, activin/inhibin activity,
CC chemotactic/chemokinetic activity, haemostatic and thrombolytic
CC activity, receptor/ligand activity, anti-inflammatory activity,
CC cadherin/tumour invasion suppressor activity, tumour inhibition
CC activity. The polynucleotide may also be useful for gene therapy.
XX
SQ Sequence 447 BP; 120 A; 107 C; 95 G; 125 T; 0 other;

Query Match 8.6%; Score 19; DB 20; Length 447;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 178 caaggggtaggagaaagag 196
||||||||||||||||||||
Db 396 caaggggtaggagaaagag 414

RESULT 25
AAQ06308/c
ID AAQ06308 standard; DNA; 710 BP.
XX
AC AAQ06308;
XX
DT 29-JAN-1991 (first entry)
XX
DE Sequence of DNA fragment F7 of the human IFN-gamma receptor.
XX
KW IFN-gamma receptor; autoimmune disease; multiple sclerosis;
KW hypersensitivity; ds.
XX
OS Homo sapiens.
XX
PN EP393502-A.
XX
PD 24-OCT-1990.
XX
PF 11-APR-1990; 90EP-0106992.
XX
PR 19-APR-1989; 89EP-0810295.
XX
PA (HOFF) HOFFMANN-LA ROCHE AG.
XX
PI Fountoulakis M, Garotta G, Stuber D;
XX
DR WPI; 1990-322042/43.
DR P-PSDB; AAR07472.
XX
PT Soluble interferon-gamma receptors - for treating auto-immune
PT diseases, chronic inflammations, etc.
XX
PS Disclosure; Fig 18; 174pp; English.
XX
CC IFN-gamma is a therapeutically active agent in the treatment
CC of autoimmune disease, allograft transplant rejections, multiple
CC sclerosis, chronic inflammations and delayed hypersensitivity. It is
CC also useful in identifying IFN-gamma agonists and antagonists.
CC See also AAQ06301.
XX
SQ Sequence 710 BP; 206 A; 147 C; 166 G; 191 T; 0 other;

Query Match 8.6%; Score 19; DB 11; Length 710;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 178 caaggggtaggagaaagag 196
||||||||||||||||||||
Db 94 CAAGGGGTAGGAGAAAGAG 76

RESULT 26
AAQ73229/c
ID AAQ73229 standard; cDNA; 735 BP.
XX
AC AAQ73229;
XX
DT 11-APR-1995 (first entry)
XX
DE Soluble human interferon gamma receptor coding sequence.
XX
KW Interferon; gamma; IFN; receptor; immunoglobulin; constant domain;
KW light chain; heavy chain; Ig; chimeric protein; fusion protein;
KW autoimmune disease; chronic inflammation; allotransplant; rejection;
KW multiple sclerosis; fulminant hepatitis; neurological disease; AIDS;
KW poliovirus; Lyme disease; scepticaemia; treatment; therapy;
KW delayed type hypersensitivity; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..735
FT /*tag= a
FT /product= Soluble interferon gamma receptor.
FT sig_peptide 1..51
FT /*tag= b
FT mat_peptide 52..735
FT /*tag= c
XX
PN EP614981-A.
XX
PD 14-SEP-1994.
XX
PF 18-FEB-1994; 94EP-0102452.
XX
PR 05-MAR-1993; 93EP-0810170.
XX
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX
PI Dembic Z, Garotta G, Gentz R;
XX
DR WPI; 1994-281208/35.
DR P-PSDB; AAR62023.
XX
PT Chimeric human interferon-gamma receptor/immunoglobulin proteins
PT - used to inhibit binding of interferon-gamma to its specific
PT receptor in the treatment of illnesses
XX
PS Disclosure; Figure 1; 29pp; English.

```
XX
CC   The soluble form of the interferon (IFN) gamma receptor comprises
CC   the whole extracellular domain of the natural receptor from the N-
CC   terminus to the transmembrane region, lacks the cytoplasmic and
CC   transmembrane domains of the natural receptor and specifically binds
CC   IFN-gamma.  The sequence encoding the soluble IFN-gamma receptor can
CC   be used in constructs encoding chimeric proteins where the other
CC   component of the chimeric protein is part or all of the constant
CC   domain of a human immunoglobulin heavy or light chain.  The
CC   recombinant proteins can be used to inhibit IFN-gamma binding to its
CC   specific receptor.  They can be used for the treatment of
CC   illnesses, especially autoimmune diseases, chronic inflammation,
CC   delayed type hypersensitivity, allotransplant rejections, multiple
CC   sclerosis, fulminant hepatitis, inflammatory neurological diseases
CC   and neurological complications of AIDS, poliovirus infections, Lyme
CC   disease and scepticaemia.  The presence of the immunoglobulin
CC   component in the chimeric protein increases the proteins half life in
CC   vivo.
XX
SQ   Sequence 735 BP; 223 A; 135 C; 170 G; 207 T; 0 other;


  Query Match             8.6%;  Score 19;  DB 15;  Length 735;
  Best Local Similarity   100.0%;  Pred. No. 5;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    178 caaggggtaggagaaagag 196
          ||||||||||||||||||||
Db     28 CAAGGGGTAGGAGAAAGAG 10


RESULT  27
AAQ06301/c
ID   AAQ06301 standard; DNA; 2184 BP.
XX
AC   AAQ06301;
XX
DT   29-JAN-1991  (first entry)
XX
DE   Sequence encoding SacI/Asp718I fragment of plasmid pBABLUE carrying
DE   human interferon-gamma receptor gene.
XX
KW   IFN-gamma receptor; autoimmune disease; multiple sclerosis;
KW   hypersensitivity; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             85..1551
FT                   /*tag=  a
XX
PN   EP393502-A.
XX
PD   24-OCT-1990.
XX
PF   11-APR-1990;   90EP-0106992.
XX
PR   19-APR-1989;   89EP-0810295.
XX
PA   (HOFF ) HOFFMANN-LA ROCHE AG.
XX
PI   Fountoulakis M,  Garotta G,  Stuber D;
XX
DR   WPI; 1990-322042/43.
DR   P-PSDB; AAR07469.
XX
PT   Soluble interferon-gamma receptors - for treating auto-immune
PT   diseases, chronic inflammations, etc.
XX
PS   Disclosure; Fig 1; 174pp; English.
XX
CC   Sequence may be used to transform prokaryotic or mammalian host
CC   cells via an expression vector, allowing production of the IFN-gamma
CC   receptor in pure form.
CC   The gene product is a therapeutically active agent in the treatment
CC   of autoimmune disease, allograft transplant rejections, multiple
CC   sclerosis, chronic inflammations and delayed hypersensitivity. It is
CC   also useful in identifying IFN-gamma agonists and antagonists.
XX
SQ   Sequence 2184 BP; 688 A; 413 C; 451 G; 632 T; 0 other;


  Query Match             8.6%;  Score 19;  DB 11;  Length 2184;
  Best Local Similarity   100.0%;  Pred. No. 4.9;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    178 caaggggtaggagaaagag 196
          ||||||||||||||||||||
Db    112 CAAGGGGTAGGAGAAAGAG 94


RESULT  28
AAC13639/c
ID   AAC13639 standard; cDNA; 367 BP.
XX
AC   AAC13639;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 17714.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-0200610.
XX
PR   26-FEB-1999;   99US-0122487.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS   Claim 1; SEQ ID 17714; 71pp + CD-ROM; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC   identified within the present sequence. The 5' ESTs were prepared from
CC   total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC   sequences usually correspond mainly to the 3' untranslated region (UTR)
CC   of the mRNA because they are often obtained from oligo-dT primed cDNA
CC   libraries. Such ESTs are not well suited for isolating cDNA sequences
CC   derived from the 5' ends of mRNAs and even in those cases where longer
CC   cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC   5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC   used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC   in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC   They are used to obtain upstream regulatory sequences and to design
CC   expression and secretion vectors.
XX
SQ   Sequence 367 BP; 88 A; 74 C; 62 G; 143 T; 0 other;


  Query Match             8.1%;  Score 18;  DB 21;  Length 367;
  Best Local Similarity   100.0%;  Pred. No. 16;
```

```
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     202 aacacaaaagtggaaaac 219
           ||||||||||||||||||
Db     104 AACACAAAAGTGGAAAAC 87


RESULT  29
AAF58344/c
ID   AAF58344 standard; cDNA; 710 BP.
XX
AC   AAF58344;
XX
DT   19-APR-2001  (first entry)
XX
DE   Human GTP-binding associated protein #44 coding sequence.
XX
KW   Human; guanosine triphosphate binding associated protein; GTP; GBAP;
KW   inflammation; AIDS; Addison's disease; anaemia; arteriosclerosis; asthma;
KW   autoimmune disorder; hepatitis; multiple sclerosis; cancer; diabetes;
KW   osteoporosis; psoriasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200105970-A2.
XX
PD   25-JAN-2001.
XX
PF   19-JUL-2000; 2000WO-US19698.
XX
PR   19-JUL-1999;   99US-0144595.
PR   23-AUG-1999;   99US-0150460.
PR   15-OCT-1999;   99US-0159849.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Yue H,  Tang YT,  Bandman O,  Hillman JL,  Lal P,  Au-Young J;
PI   Reddy R,  Yang J,  Baughn MR,  Lu DAM,  Azimzai Y,  Patterson C;
XX
DR   WPI; 2001-091972/10.
DR   P-PSDB; AAB68544.
XX
PT   New guanosine triphosphate-binding associated proteins (GBAP) and their
PT   encoding nucleic acids, useful for treating and/or diagnosing diseases
PT   associated with GBAP expression, such as cancer, diabetes and asthma -
XX
PS   Claim 5; Page 216; 233pp; English.
XX
CC   The present invention relates to novel human guanosine triphosphate
CC   (GTP)-binding associated proteins (GBAPs; AAB68501-AAB68566) and their
CC   coding sequences (AAF58301-AAF58366). The proteins and coding sequences
CC   of the present invention are useful for treating a variety of disorders
CC   including inflammation, AIDS, Addison's disease, anaemia,
CC   arteriosclerosis, asthma, autoimmune disorders, Grave's disease,
CC   hepatitis, multiple sclerosis, cancer, diabetes, osteoporosis and
CC   psoriasis.
XX
SQ   Sequence 710 BP; 145 A; 290 C; 153 G; 122 T; 0 other;


  Query Match             8.1%;  Score 18;  DB 22;  Length 710;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      14 agagtgagtggggccggg 31
           ||||||||||||||||||
Db     555 AGAGTGAGTGGGGCCGGG 538


RESULT  30
AAF15226/c
ID   AAF15226 standard; cDNA; 758 BP.
XX
AC   AAF15226;
XX
DT   13-MAR-2001  (first entry)
XX
DE   Trichoderma reesei EST SEQ ID NO:7749.
XX
KW   Multiple gene expression; filamentous fungal cell; EST;
KW   expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KW   Aspergillus oryzae; Trichoderma reesei; identification; recombination;
KW   culture condition; environmental stress; spore morphogenesis;
KW   metabolic pathway engineering; catabolic pathway engineering; ss.
XX
OS   Trichoderma reesei.
XX
PN   WO200056762-A2.
XX
PD   28-SEP-2000.
XX
PF   22-MAR-2000; 2000WO-US07781.
XX
PR   22-MAR-1999;   99US-0273623.
XX
PA   (NOVO ) NOVO NORDISK BIOTECH INC.
PA   (NOVO ) NOVO NORDISK AS.
XX
PI   Berka RM,  Rey MW,  Shuster JR,  Kauppinen S,  Clausen IG,  Olsen PB;
XX
DR   WPI; 2000-594572/56.
XX
PT   Monitoring differential expression of genes in filamentous fungal cells
PT   uses fluorescence-labeled nucleic acids isolated from the cells and a
PT   substrate of expressed sequence tags -
XX
PS   Claim 89; Page 3125; 3161pp; English.
XX
CC   The present invention describes a method for monitoring differential
CC   expression of genes in a first filamentous fungal (FF) cell relative to
CC   expression of the same genes in one or more second filamentous fungal
CC   cells. The method uses fluorescence-labeled nucleic acids isolated from
CC   the FF cells and a substrate of expressed sequence tags (EST). The ESTs
CC   are used in the methods for monitoring differential expression of genes
CC   in a first filamentous fungal (FF) cell relative to expression of the
CC   same genes in one or more second filamentous fungal cells. Monitoring
CC   the global expression of genes from FF cells allows the production
CC   potential of the microorganisms to be improved. New genes may be
CC   discovered, possible functions of unknown open reading frames can be
CC   identified and gene copy number variation and stability can be
CC   monitored. The expression of genes can be used to study how FF cells
CC   adapt to changes in culture conditions, environmental stress, spore
CC   morphogenesis, recombination, metabolic or catabolic pathway
CC   engineering. Using ESTs provides several advantages over genomic or
CC   random cDNA clones including elimination of redundancy as one spot on an
CC   array equals one gene or open reading frame, and organisation of the
CC   microarrays based on function of the gene products to facilitate
CC   analysis of the results. AAF07478 to AAF11247 represents ESTs from
CC   Fusarium venenatum; AAF11248 to AAF11853 represents ESTs from Aspergillus
CC   niger; AAF11854 to AAF14878 represents ESTs from Aspergillus oryzae; and
CC   AAF14879 to AAF15337 represents ESTs from Trichoderma reesei, which are
CC   all specifically claimed in the present invention.
XX
SQ   Sequence 758 BP; 184 A; 168 C; 170 G; 199 T; 37 other;


  Query Match             8.1%;  Score 18;  DB 21;  Length 758;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     131 ttgttttttccccggttc 148
           ||||||||||||||||||
Db     487 TTGTTTTTTCCCCGGTTC 470


RESULT  31
```

```
AAS68374
ID   AAS68374 standard; cDNA; 1166 BP.
XX
AC   AAS68374;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #4178.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG04187.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 4178; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1166 BP; 283 A; 299 C; 293 G; 278 T; 13 other;

  Query Match             8.1%;  Score 18;  DB 23;  Length 1166;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      63 gggctgcggcagggcagg 80
           ||||||||||||||||||
Db     182 gggctgcggcagggcagg 199


RESULT  32
AAS72684/c
```

```
ID   AAS72684 standard; cDNA; 1287 BP.
XX
AC   AAS72684;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #8488.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG08497.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 8488; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1287 BP; 237 A; 461 C; 391 G; 198 T; 0 other;

  Query Match             8.1%;  Score 18;  DB 23;  Length 1287;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      64 ggctgcggcagggcaggg 81
           ||||||||||||||||||
Db     734 GGCTGCGGCAGGGCAGGG 717


RESULT  33
AAQ66990
ID   AAQ66990 standard; DNA; 1656 BP.
```

```
XX
AC   AAQ66990;
XX
DT   18-MAR-1995  (first entry)
XX
DE   DF3 enhancer.
XX
KW   Tissue-specific expression; epithelium; carcinoma; metabolic;
KW   disorders; phenylketonuria; cystic fibrosis; haemophilia; emphysema;
KW   vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO9416057-A.
XX
PD   21-JUL-1994.
XX
PF   08-DEC-1993;   93WO-US11915.
XX
PR   31-DEC-1992;   92US-0999742.
XX
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Abe M,  Kufe D;
XX
DR   WPI; 1994-249205/30.
XX
PT   Isolated DF3 enhancer DNA sequence - is used to enhance
PT   epithelium tissue-specific expression for treating carcinomas or
PT   metabolic disorders
XX
PS   Claim 1; Fig 7; 43pp; English.
XX
CC   The sequence represents the DF3 gene 5' flanking sequence. The DF3
CC   enhancer can be linked to a heterologous coding sequence to
CC   increase tissue-specific expression of the coding sequence in
CC   the epithelium.  These prods. can be used to treat carcinomas or
CC   metabolic disorders such as phenylketonuria, cystic fibrosis,
CC   haemophilia or emphysema.  They can also be used as vaccines
CC   against infectious organisms or can confer resistance to
CC   the effects of a chemotherapeutic drug.
CC   See also AAQ66988-95.
XX
SQ   Sequence 1656 BP; 275 A; 556 C; 478 G; 347 T; 0 other;


  Query Match             8.1%;  Score 18;  DB 15;  Length 1656;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 gacccttctctcccgggc 66
           ||||||||||||||||||
Db     773 gacccttctctcccgggc 790


RESULT  34
AAQ66988
ID   AAQ66988 standard; DNA; 1725 BP.
XX
AC   AAQ66988;
XX
DT   18-MAR-1995  (first entry)
XX
DE   5' flanking region oof the DF3 gene.
XX
KW   Tissue-specific expression; epithelium; carcinoma; metabolic;
KW   disorders; phenylketonuria; cystic fibrosis; haemophilia; emphysema;
KW   vaccine; enhancer; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   misc_RNA        1657
FT                   /*tag=  a
FT                   /note= "transcription start site"
FT   TATA_signal     1627
FT                   /*tag=  b
XX
PN   WO9416057-A.
XX
PD   21-JUL-1994.
XX
PF   08-DEC-1993;   93WO-US11915.
XX
PR   31-DEC-1992;   92US-0999742.
XX
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Abe M,  Kufe D;
XX
DR   WPI; 1994-249205/30.
XX
PT   Isolated DF3 enhancer DNA sequence - is used to enhance
PT   epithelium tissue-specific expression for treating carcinomas or
PT   metabolic disorders
XX
PS   Disclosure; Fig 1A; 43pp; English.
XX
CC   The sequence represents the DF3 gene 5' flanking sequence. The DF3
CC   enhancer can be linked to a heterologous coding sequence to
CC   increase tissue-specific expression of the coding sequence in
CC   the epithelium.  These prods. can be used to treat carcinomas or
CC   metabolic disorders such as phenylketonuria, cystic fibrosis,
CC   haemophilia or emphysema.  They can also be used as vaccines
CC   against infectious organisms or can confer resistance to
CC   the effects of a chemotherapeutic drug.
CC   See also AAQ66989-95.
XX
SQ   Sequence 1725 BP; 288 A; 588 C; 488 G; 361 T; 0 other;


  Query Match             8.1%;  Score 18;  DB 15;  Length 1725;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 gacccttctctcccgggc 66
           ||||||||||||||||||
Db     773 gacccttctctcccgggc 790


RESULT  35
ABA20286/c
ID   ABA20286 standard; DNA; 2484 BP.
XX
AC   ABA20286;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Human nervous system related polynucleotide SEQ ID NO 12617.
XX
KW   Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW   immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW   antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW   antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW   antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW   antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS   Homo sapiens.
XX
PN   WO200159063-A2.
XX
PD   16-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01334.
XX
```

PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-541565/60.
XX
PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system
PT cancers and metastases -

```
XX
PS   Disclosure; SEQ ID NO 12617; 1701pp + Sequence Listing; English.
XX
CC   The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC   (ABB14678-ABB18001) useful for preventing, treating or ameliorating
CC   medical conditions e.g. by protein or gene therapy. The genes are
CC   isolated from a range of human tissues disclosed in the specification.
CC   The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC   in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC   and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC   marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC   (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC   haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC   disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC   colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC   (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC   epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC   and parasitic infections.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2484 BP; 440 A; 906 C; 734 G; 404 T; 0 other;


  Query Match              8.1%;  Score 18;  DB 22;  Length 2484;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      15 gagtgagtggggccggga 32
           ||||||||||||||||||
Db    2406 GAGTGAGTGGGGCCGGGA 2389


RESULT  36
ABA20287/c
ID   ABA20287 standard; DNA; 2484 BP.
XX
AC   ABA20287;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Human nervous system related polynucleotide SEQ ID NO 12618.
XX
KW   Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW   immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW   antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW   antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW   antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW   antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS   Homo sapiens.
XX
PN   WO200159063-A2.
XX
PD   16-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01334.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
```

PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-541565/60.
XX
PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system
PT cancers and metastases -
XX
PS Disclosure; SEQ ID NO 12618; 1701pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (ABB14678-ABB18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2484 BP; 440 A; 906 C; 734 G; 404 T; 0 other;

Query Match 8.1%; Score 18; DB 22; Length 2484;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 gagtgagtggggccggga 32
||||||||||||||||||
Db 2406 GAGTGAGTGGGGCCGGGA 2389

RESULT 37
AAZ86925/c
ID AAZ86925 standard; cDNA; 2779 BP.
XX
AC AAZ86925;
XX
DT 04-MAY-2000 (first entry)
XX
DE Netrin-2 coding sequence.
XX
KW Netrin-1; netrin-2; binding assay; neural axon outgrowth; diagnosis;
KW netrin-specific antibody; neurological disease; therapy; ss.
XX
OS Unidentified.
XX
PN US6017714-A.
XX
PD 25-JAN-2000.
XX
PF 07-JUN-1995; 95US-0482677.
XX
PR 12-NOV-1993; 93US-0152019.
XX
PA (REGC) UNIV CALIFORNIA.
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Serafini T, Kennedy T, Tessier-Lavigne M, Jessell T, Dodd J;
PI Placzek M;
XX
DR WPI; 2000-136674/12.
DR P-PSDB; AAY76838.
XX
PT New binding assay for identifying agents which specifically bind to
PT netrin proteins which can be used as diagnostic tools for neurological
PT diseases -
XX
PS Disclosure; Column 17-20; 23pp; English.
XX
CC This sequence encodes a netrin protein. The invention relates to a
CC binding assay using a netrin or portion of a netrin protein to identify
CC an agent which specifically binds a netrin. The assay comprises:
CC (1) contacting a prospective agent with a netrin which modulates axon
CC outgrowth or guidance or elicits a netrin-specific antibody; and
CC (2) determining if the agent specifically binds netrin. Vertebrate netrin
CC proteins are involved in modulating neural axon outgrowth and identifying
CC agents which bind or modulate netrin function allows identification of
CC regulators of axon outgrowth and orientation which can be used as
CC diagnostic reagents for neurological disease and for the development of
CC neurological disease therapy. Natural and synthetic chemical libraries
CC can be screened using these methods to identify reagents suitable for use
CC in human and animal clinical trials and pharmacological agents or lead
CC compounds for agents capable of mimicking or modulating netrin function

```
CC   can also be identified. The methods are cost-effective, amenable to
CC   automation and can provide high throughput screening.
XX
SQ   Sequence 2779 BP; 565 A; 897 C; 797 G; 520 T; 0 other;


  Query Match             8.1%;  Score 18;  DB 21;  Length 2779;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      20 agtggggccgggacccgc 37
           ||||||||||||||||||
Db    1782 AGTGGGGCCGGGACCCGC 1765


RESULT  38
AAQ92367/c
ID   AAQ92367 standard; cDNA; 2783 BP.
XX
AC   AAQ92367;
XX
DT   24-DEC-1995  (first entry)
XX
DE   Chick p75 cDNA.
XX
KW   Neural axon out-growth modulator; epidermal growth factor; EGF;
KW   netrin-2; p75; neurodegenerative disease; transgenic animal;
KW   gene therapy; ss.
XX
OS   Gallus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             4..1749
FT                   /*tag=  a
XX
PN   WO9513367-A1.
XX
PD   18-MAY-1995.
XX
PF   08-NOV-1994;   94WO-US12913.
XX
PR   12-NOV-1993;   93US-0152019.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYCO ) UNIV COLUMBIA NEW YORK.
XX
PI   Dodd J,  Jessell T,  Kennedy T,  Placzek M,  Serafini T;
PI   Tessier-Lavigne M;
XX
DR   WPI; 1995-194086/25.
DR   P-PSDB; AAR74187.
XX
PT   Neural axon out-growth modulators derived from EGF-like repeats of
PT   netrin 1 or netrin 2 - comprise peptide(s) capable of selectively
PT   increasing spinal axon out-growth or directing axon orientation
XX
PS   Disclosure; Page 46-48; 58pp; English.
XX
CC   An E10 chick brain cDNA library was screened with probes based on
CC   netrin-1 (p78) or netrin-2 (p75) sequences to isolate chick p78
CC   and p75 partial cDNA clones.  Full-length clones (given in
CC   AAQ92366-67, respectively) were subsequently obtd. by 3'RACE.  cDNA
CC   is expressed e.g. in COS or insect cells for recombinant p78 and
CC   p75 prodn., used to breed transgenic animals, or for gene therapy.
XX
SQ   Sequence 2783 BP; 565 A; 897 C; 797 G; 520 T; 4 other;


  Query Match             8.1%;  Score 18;  DB 16;  Length 2783;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      20 agtggggccgggacccgc 37
           ||||||||||||||||||
Db    1783 AGTGGGGCCGGGACCCGC 1766


RESULT  39
AAZ32011/c
ID   AAZ32011 standard; DNA; 9248 BP.
XX
AC   AAZ32011;
XX
DT   10-JAN-2000  (first entry)
XX
DE   Human METH1 related EST AB001735.
XX
KW   Human; METH1; METH2; anti-angiogenic; metalloprotease thrombospondin;
KW   cancer; diagnosis; hyperproliferative disorder; autoimmune disease;
KW   angiogenesis inhibitor; abnormal wound healing; inflammation;
KW   rheumatoid arthritis; psoriasis; endometrial bleeding disorder;
KW   diabetic retinopathy; macula degeneration; haemangioma; detection;
KW   arterial-venous malformation; immune deficiency; ss.
XX
OS   Homo sapiens.
XX
PN   WO9937660-A1.
XX
PD   29-JUL-1999.
XX
PF   22-JAN-1999;   99WO-US01313.
XX
PR   23-JAN-1998;   98US-0072298.
PR   28-AUG-1998;   98US-0098539.
XX
PA   (IRUE/) IRUELA-ARISPE L.
PA   (HAST/) HASTINGS G A.
PA   (RUBE/) RUBEN S M.
XX
PI   Iruela-Arispe L,  Hastings GA,  Ruben SM;
XX
DR   WPI; 1999-590684/50.
XX
PT   New isolated metalloprotease thrombospondin polypeptides, useful for
PT   treating hyperproliferative disorders, cancers or autoimmune disorders
PT   -
XX
PS   Disclosure; Page 246-252; 457pp; English.
XX
CC   AAZ32000 and AAZ32001 encode, and AAY49501 and AAY49502 represent, human
CC   metalloprotease thrombospondin (METH) proteins METH1 and METH2
CC   respectively. METH1 and METH2 have been found to be potent inhibitors of
CC   angiogenesis both in vitro and in vivo. They can be used for treating
CC   cancer and other disorders related to angiogenesis including abnormal
CC   wound healing, inflammation, rheumatoid arthritis, psoriasis,
CC   endometrial bleeding disorders, diabetic retinopathy, some forms of
CC   macula degeneration, haemangiomas, and arterial-venous malformations.
CC   They may be useful in treating deficiencies or disorders of the immune
CC   system, by activating or inhibiting the proliferation, differentiation,
CC   or mobilisation (chemotaxis) of immune cells. The etiology of these
CC   immune deficiencies or disorders may be genetic, somatic, such as
CC   cancer or some autoimmune disorders, acquired (e.g. by chemotherapy or
CC   toxins), or infectious. They can also be used to treat inflammatory
CC   conditions, both chronic and acute conditions. The products can also be
CC   used for detection and diagnosis. AAZ32002 to AAZ32080, and AAY49503 to
CC   AAY49511 represent sequences given in the exemplification of the present
CC   invention.
XX
SQ   Sequence 9248 BP; 2475 A; 2123 C; 2207 G; 2443 T; 0 other;


  Query Match             8.1%;  Score 18;  DB 20;  Length 9248;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      74 gggcagggcggggagctc 91
```

```
                  ||||||||||||||||||
Db    2223 GGGCAGGGCGGGGAGCTC 2206


RESULT  40
AAC90068/c
ID   AAC90068 standard; DNA; 9248 BP.
XX
AC   AAC90068;
XX
DT   19-MAR-2001  (first entry)
XX
DE   AB001735 cDNA clone.
XX
KW   METH; metalloprotease; thrombospondin; angiogenesis inhibition;
KW   cancer therapy; benign tumour; ocular angiogenic disease;
KW   rheumatoid arthritis; psoriasis; wound healing; endometriosis;
KW   vasculogenesis; granulation; hypertrophic scar; nonunion fracture;
KW   scleroderma, trachoma; vascular adhesion; myocardial angiogenesis;
KW   coronary collateral; cerebral collateral; arteriovenous malformation;
KW   ischaemic limb angiogenesis; Osler-Webber syndrome; wound granulation;
KW   plaque neovascularisation; telangiectasia; haemophiliac joint; EST;
KW   angiofibroma; fibromuscular dysplasia; expressed sequence tag;
KW   Crohn's disease; atherosclerosis; birth control; ss.
XX
OS   Unidentified.
XX
PN   WO200071577-A1.
XX
PD   30-NOV-2000.
XX
PF   25-MAY-2000; 2000WO-US14462.
XX
PR   25-MAY-1999;   99US-0318208.
PR   20-JUL-1999;   99US-0144882.
PR   10-AUG-1999;   99US-0147823.
PR   13-AUG-1999;   99US-0373658.
PR   22-DEC-1999;   99US-0171503.
PR   22-FEB-2000; 2000US-0183792.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
PA   (IRUE/) IRUELA-ARISPE L.
PA   (HAST/) HASTINGS G A.
PA   (RUBE/) RUBEN S M.
PA   (JONA/) JONAK Z L.
PA   (TRUL/) TRULLI S H.
PA   (FORN/) FORNWALD J A.
PA   (TERR/) TERRETT J A.
XX
PI   Iruela-Arispe L,  Hastings GA,  Ruben SM,  Jonak ZL,  Trulli SH;
PI   Fornwald JA,  Terrett JA;
XX
DR   WPI; 2001-025136/03.
XX
PT   METH1 and METH2 polynucleotides and encoded polypeptides, used to
PT   inhibit angiogensis in the treatment of disorders such as cancer,
PT   rheumatoid arthritis and psoriasis -
XX
PS   Claim 7; Pages 546-552; 768pp; English.
XX
CC   The present invention relates to human METH1 and METH2, (ME for
CC   metalloprotease and TH for thrombospondin; see AAB50002 and AAB50003).
CC   The present sequence is an expressed sequence tag (EST) for METH. METH
CC   can be used for inhibiting angiogenesis in an individual, and for
CC   treating cancer, benign tumours, an ocular angiogenic disease,
CC   rheumatoid arthritis, psoriasis, delayed wound healing, endometriosis,
CC   vasculogenesis, granulations, hypertrophic scars, nonunion fractures,
CC   scleroderma, trachoma, vascular adhesions, myocardial angiogenesis,
CC   coronary collaterals, cerebral collaterals, arteriovenous malformations,
CC   ischaemic limb angiogenesis, Osler-Webber syndrome, plaque
CC   neovascularisation, telangiectasia, haemophiliac joints, angiofibroma,
CC   fibromuscular dysplasia, wound granulation, Crohn's disease or
CC   atherosclerosis. METH can also be used in birth control. METH can also
CC   be used in diagnostic methods for the prognosis of cancer.
XX
SQ   Sequence 9248 BP; 2475 A; 2123 C; 2207 G; 2443 T; 0 other;


  Query Match             8.1%;  Score 18;  DB 22;  Length 9248;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      74 gggcagggcggggagctc 91
           ||||||||||||||||||
Db    2223 GGGCAGGGCGGGGAGCTC 2206


RESULT  41
ABL32987
ID   ABL32987 standard; DNA; 17294 BP.
XX
AC   ABL32987;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human immune system associated gene SEQ ID NO: 960.
XX
KW   Human; immune system disease; cytosine methylation; antiasthmatic;
KW   antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW   neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW   antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW   antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW   acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW   neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW   gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200200928-A2.
XX
PD   03-JAN-2002.
XX
PF   02-JUL-2001; 2001WO-EP07537.
XX
PR   30-JUN-2000; 2000DE-1032529.
PR   01-SEP-2000; 2000DE-1043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-130909/17.
XX
PT   Nucleic acid comprising fragment of chemically modified gene, useful
PT   for diagnosis and treatment of diseases associated with abnormal
PT   cytosine methylation  -
XX
PS   Claim 1; SEQ ID NO 960; 32pp + Sequence Listing; German.
XX
CC   The present invention provides a number of human immune system associated
CC   genes which are modified by the methylation of cytosines. The sequences
CC   can be used in the diagnosis and treatment of immune system disorders,
CC   including eye diseases such as retinopathy, neovascular glaucoma and
CC   macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC   leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC   rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC   diseases. The present sequence is a gene of the invention.
XX
SQ   Sequence 17294 BP; 5081 A; 203 C; 3600 G; 8410 T; 0 other;


  Query Match             8.1%;  Score 18;  DB 24;  Length 17294;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy    179 aaggggtaggagaaagag 196
          ||||||||||||||||||
Db  12225 aaggggtaggagaaagag 12242


RESULT  42
AAK70011
ID   AAK70011 standard; DNA; 21436 BP.
XX
AC   AAK70011;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:24823.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01354.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
```

```
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-483426/52.
XX
PT   Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating cancers and
PT   metastasis -
XX
PS   Disclosure; SEQ ID NO 24823; 3071pp + Sequence Listing; English.
XX
CC   AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production. (I)
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting
CC   the nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention.
XX
SQ   Sequence 21436 BP; 5267 A; 4950 C; 5418 G; 5801 T; 0 other;


  Query Match             8.1%;  Score 18;  DB 22;  Length 21436;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      64 ggctgcggcagggcaggg 81
           ||||||||||||||||||
Db   18126 ggctgcggcagggcaggg 18143


RESULT  43
AAK79799
ID   AAK79799 standard; DNA; 21436 BP.
XX
AC   AAK79799;
XX
DT   07-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:34611.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01354.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
```

PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Disclosure; SEQ ID NO 34611; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 21436 BP; 5267 A; 4950 C; 5418 G; 5801 T; 0 other;

Query Match 8.1%; Score 18; DB 22; Length 21436;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 64 ggctgcggcagggcaggg 81
||||||||||||||||||
Db 18126 ggctgcggcagggcaggg 18143

RESULT 44
AAF92943
ID AAF92943 standard; DNA; 17 BP.
XX
AC AAF92943;
XX
DT 17-MAY-2001 (first entry)
XX
DE Wild type sequence for ABC1 polymorphic site #16.
XX
KW High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ds.
XX
OS Homo sapiens.
XX
PN WO200115676-A2.
XX
PD 08-MAR-2001.
XX
PF 01-SEP-2000; 2000WO-IB01492.
XX
PR 01-SEP-1999; 99US-0151977.
PR 15-MAR-2000; 2000US-0526193.
PR 23-JUN-2000; 2000US-0213958.
XX
PA (UYBR-) UNIV BRITISH COLUMBIA.
PA (XENO-) XENON GENETICS INC.
XX
PI Hayden MR, Brooks-Wilson AR, Pimstone SN, Clee SM;
XX
DR WPI; 2001-244356/25.
XX
PT Treating a lower than normal high density lipoprotein-cholesterol

PT (HDL-C) level, a higher than normal triglyceride level, or a
PT cardiovascular disease, by administering a compound that modulates LXR-
PT or RXR-mediated transcriptional activity -
XX
PS Disclosure; Fig 4; 317pp; English.
XX
CC The present invention relates to a method for treating a patient
CC diagnosed as having a lower than normal high density
CC lipoprotein-cholesterol (HDL-C) level, a higher than normal
CC triglyceride level, or a cardiovascular disease, involving
CC administering a compound that modulates LXR- or RXR-mediated
CC transcriptional activity or ABC1 expression or activity.
CC The LXR gene product may be used in an assay to identify
CC compounds useful for the treatment of a disease or condition selected a
CC lower than normal HDL cholesterol level, a higher than normal
CC triglyceride level, and a cardiovascular disease.
XX
SQ Sequence 17 BP; 3 A; 8 C; 6 G; 0 U; 0 other;

Query Match 7.7%; Score 17; DB 22; Length 17;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 27 ccgggacccgcagagcc 43
||||||||||||||||| -
Db 1 ccgggacccgcagagcc 17

RESULT 45
AAK90657/c
ID AAK90657 standard; DNA; 125 BP.
XX
AC AAK90657;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human digestive system antigen genomic sequence SEQ ID NO: 4233.
XX
KW Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum; ds.
XX
OS Homo sapiens.
XX
PN WO200155314-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01324.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.

PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-502630/55.
XX
PT Polynucleotides encoding digestive system antigens, useful for
PT diagnosing, treating, preventing and/or prognozing disorders of the
PT digestive system, particularly cancer and cancer metastases -
XX
PS Disclosure; SEQ ID NO 4233; 986pp; English.
XX
CC The present invention provides the protein and coding sequences of a
CC number of human digestive system antigens. These can be used in the
CC diagnosis, treatment and prevention of digestive system disorders,
CC including cancer, Meckel's diverticulum, bacterial or parasitic
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC ulcerative colitis. The present sequence is a genomic DNA fragment
CC encoding a digestive system antigen of the invention.
XX
SQ Sequence 125 BP; 37 A; 13 C; 23 G; 52 T; 0 other;

Query Match 7.7%; Score 17; DB 22; Length 125;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 204 cacaaaagtggaaaaca 220
||||||||||||||||
Db 63 CACAAAAGTGGAAAACA 47

Search completed: September 20, 2002, 06:08:05
Job time: 10369 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 06:15:23 ; Search time 111.56 Seconds
(without alignments)
486.599 Million cell updates/sec

Title: US-09-846-456-4
Perfect score: 221
Sequence: 1 gtaattgcgagcgagagtga.....aacacaaaagtggaaaacag 221

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	18	8.1	1656	1	US-08-324-465-2	Sequence 2, Appli
	2	18	8.1	1656	2	US-08-465-981-2	Sequence 2, Appli
	3	18	8.1	1656	5	PCT-US93-11915-2	Sequence 2, Appli
	4	18	8.1	1725	1	US-08-324-465-5	Sequence 5, Appli
	5	18	8.1	1725	2	US-08-465-981-5	Sequence 5, Appli
	6	18	8.1	1725	5	PCT-US93-11915-5	Sequence 5, Appli
c	7	18	8.1	2779	3	US-08-482-677-5	Sequence 5, Appli
c	8	18	8.1	2783	1	US-08-152-019A-41	Sequence 41, Appl
	9	16	7.2	544	2	US-08-890-980-17	Sequence 17, Appl
	10	16	7.2	544	3	US-08-890-979-17	Sequence 17, Appl
	11	16	7.2	544	3	US-09-032-894-17	Sequence 17, Appl
	12	16	7.2	544	4	US-09-031-626-17	Sequence 17, Appl
	13	16	7.2	1002	2	US-08-890-980-5	Sequence 5, Appli
	14	16	7.2	1002	3	US-08-890-979-5	Sequence 5, Appli
	15	16	7.2	1002	3	US-09-032-894-5	Sequence 5, Appli
	16	16	7.2	1002	3	US-09-032-894-95	Sequence 95, Appl
	17	16	7.2	1002	4	US-09-031-626-5	Sequence 5, Appli
	18	16	7.2	1002	4	US-09-031-626-95	Sequence 95, Appl
	19	16	7.2	1209	4	US-09-105-537-21	Sequence 21, Appl
	20	16	7.2	1494	1	US-08-318-905-21	Sequence 21, Appl
	21	16	7.2	1494	1	US-08-483-232-21	Sequence 21, Appl
	22	16	7.2	1494	1	US-08-483-140-21	Sequence 21, Appl
	23	16	7.2	1494	2	US-08-485-938A-21	Sequence 21, Appl
	24	16	7.2	1494	2	US-08-910-041-21	Sequence 21, Appl
	25	16	7.2	1494	3	US-09-328-474-21	Sequence 21, Appl
	26	16	7.2	1494	3	US-09-100-546-21	Sequence 21, Appl
	27	16	7.2	1494	3	US-09-010-715-21	Sequence 21, Appl
	28	16	7.2	1494	4	US-09-577-758-21	Sequence 21, Appl
	29	16	7.2	1613	2	US-08-812-204-1	Sequence 1, Appli
	30	16	7.2	1881	5	PCT-US94-09752-2	Sequence 2, Appli
	31	16	7.2	2692	1	US-07-932-454A-2	Sequence 2, Appli
c	32	16	7.2	6840	4	US-08-980-241-8	Sequence 8, Appli
	33	16	7.2	13613	4	US-09-105-537-3	Sequence 3, Appli
	34	16	7.2	15202	3	US-08-922-635-21	Sequence 21, Appl
c	35	16	7.2	16836	4	US-09-147-236-1	Sequence 1, Appli
c	36	16	7.2	16836	4	US-09-147-236-10	Sequence 10, Appl
	37	16	7.2	38506	3	US-09-320-878-19	Sequence 19, Appl
c	38	15	6.8	20	4	US-09-049-714-1	Sequence 1, Appli
c	39	15	6.8	22	4	US-08-928-465-3	Sequence 3, Appli
c	40	15	6.8	24	4	US-09-481-288-2	Sequence 2, Appli
c	41	15	6.8	43	2	US-08-850-049-108	Sequence 108, App
c	42	15	6.8	43	2	US-08-050-478-108	Sequence 108, App
c	43	15	6.8	43	4	US-09-414-117-108	Sequence 108, App
	44	15	6.8	51	3	US-09-046-247-24	Sequence 24, Appl
c	45	15	6.8	239	4	US-09-481-288-1	Sequence 1, Appli

ALIGNMENTS

```
RESULT   1
US-08-324-465-2
; Sequence 2, Application US/08324465
; Patent No. 5565334
;  GENERAL INFORMATION:
;    APPLICANT:  Kufe, Donald
;    APPLICANT:  Abe, Miyako
;    TITLE OF INVENTION:  GENE TRANSCRIPTION AND
;    TITLE OF INVENTION:  IONIZING RADIATION: METHODS
;    TITLE OF INVENTION:  AND COMPOSITIONS
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/324,465
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/999,742
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fraser, Janis K.
;      REGISTRATION NUMBER:  34,819
;      REFERENCE/DOCKET NUMBER:  00530/065001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1656
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-324-465-2


  Query Match              8.1%;  Score 18;  DB 1;  Length 1656;
  Best Local Similarity   100.0%;  Pred. No. 3.6;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      49 gacccttctctcccgggc 66
           ||||||||||||||||||
Db     773 GACCCTTCTCTCCCGGGC 790


RESULT   2
US-08-465-981-2
; Sequence 2, Application US/08465981
; Patent No. 5874415
;  GENERAL INFORMATION:
;    APPLICANT:  Kufe, Donald
;    APPLICANT:  Abe, Miyako
;    TITLE OF INVENTION:  ENHANCER SEQUENCE FOR MODULATING
;    TITLE OF INVENTION:  EXPRESSION IN EPITHELIAL CELLS
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/465,981
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/324,465
;      FILING DATE:  October 17, 1994
;      APPLICATION NUMBER:  07/999,742
;      FILING DATE:  December 31, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fraser, Janis K.
;      REGISTRATION NUMBER:  34,819
;      REFERENCE/DOCKET NUMBER:  00530/065002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1656
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-465-981-2


  Query Match             8.1%;  Score 18;  DB 2;  Length 1656;
  Best Local Similarity   100.0%;  Pred. No. 3.6;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 gacccttctctcccgggc 66
           ||||||||||||||||||
Db     773 GACCCTTCTCTCCCGGGC 790


RESULT   3
PCT-US93-11915-2
; Sequence 2, Application PC/TUS9311915
;  GENERAL INFORMATION:
;    APPLICANT:  Kufe, Donald
;    APPLICANT:  Abe, Miyako
;    TITLE OF INVENTION:  ENHANCER SEQUENCE FOR  MODULATING
;    TITLE OF INVENTION:  EXPRESSION IN EPITHELIAL CELLS
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/11915
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/999,742
;      FILING DATE:  December 31, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fraser, Janis K.
;      REGISTRATION NUMBER:  34,819
;      REFERENCE/DOCKET NUMBER:  00530/065WO1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1656
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
PCT-US93-11915-2


  Query Match             8.1%;  Score 18;  DB 5;  Length 1656;
  Best Local Similarity   100.0%;  Pred. No. 3.6;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 gacccttctctcccgggc 66
           ||||||||||||||||||
Db     773 GACCCTTCTCTCCCGGGC 790


RESULT   4
US-08-324-465-5
; Sequence 5, Application US/08324465
; Patent No. 5565334
;  GENERAL INFORMATION:
;    APPLICANT:  Kufe, Donald
;    APPLICANT:  Abe, Miyako
;    TITLE OF INVENTION:  GENE TRANSCRIPTION AND
;    TITLE OF INVENTION:  IONIZING RADIATION: METHODS
;    TITLE OF INVENTION:  AND COMPOSITIONS
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/324,465
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/999,742
```

```
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fraser, Janis K.
;      REGISTRATION NUMBER:  34,819
;      REFERENCE/DOCKET NUMBER:  00530/065001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1725
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-324-465-5


  Query Match             8.1%;  Score 18;  DB 1;  Length 1725;
  Best Local Similarity   100.0%;  Pred. No. 3.6;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 gacccttctctcccgggc 66
           ||||||||||||||||||
Db     773 GACCCTTCTCTCCCGGGC 790


RESULT   5
US-08-465-981-5
; Sequence 5, Application US/08465981
; Patent No. 5874415
;  GENERAL INFORMATION:
;    APPLICANT:  Kufe, Donald
;    APPLICANT:  Abe, Miyako
;    TITLE OF INVENTION:  ENHANCER SEQUENCE FOR MODULATING
;    TITLE OF INVENTION:  EXPRESSION IN EPITHELIAL CELLS
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/465,981
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/324,465
;      FILING DATE:  October 17, 1994
;      APPLICATION NUMBER:  07/999,742
;      FILING DATE:  December 31, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fraser, Janis K.
;      REGISTRATION NUMBER:  34,819
;      REFERENCE/DOCKET NUMBER:  00530/065002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1725
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-465-981-5


  Query Match             8.1%;  Score 18;  DB 2;  Length 1725;
  Best Local Similarity   100.0%;  Pred. No. 3.6;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 gacccttctctcccgggc 66
           ||||||||||||||||||
Db     773 GACCCTTCTCTCCCGGGC 790


RESULT   6
PCT-US93-11915-5
; Sequence 5, Application PC/TUS9311915
;  GENERAL INFORMATION:
;    APPLICANT:  Kufe, Donald
;    APPLICANT:  Abe, Miyako
;    TITLE OF INVENTION:  ENHANCER SEQUENCE FOR  MODULATING
;    TITLE OF INVENTION:  EXPRESSION IN EPITHELIAL CELLS
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/11915
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/999,742
;      FILING DATE:  December 31, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fraser, Janis K.
;      REGISTRATION NUMBER:  34,819
;      REFERENCE/DOCKET NUMBER:  00530/065WO1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1725
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
PCT-US93-11915-5


  Query Match             8.1%;  Score 18;  DB 5;  Length 1725;
  Best Local Similarity   100.0%;  Pred. No. 3.6;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 gacccttctctcccgggc 66
           ||||||||||||||||||
Db     773 GACCCTTCTCTCCCGGGC 790


RESULT   7
US-08-482-677-5/c
; Sequence 5, Application US/08482677
; Patent No. 6017714
;  GENERAL INFORMATION:
;    APPLICANT:  Tessier-Lavigne, Marc
;    APPLICANT:  Serafini, Tito
;    APPLICANT:  Kennedy, Timothy
```

```
;    APPLICANT:  Placzek, Marysia
;    APPLICANT:  Jessel, Thomas
;    APPLICANT:  Dodd, Jane
;    TITLE OF INVENTION:  Netrins
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCIENCE & TECHNOLOGY LAW GROUP
;      STREET:  268 BUSH STREET, SUITE 3200
;      CITY:  SAN FRANCISCO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/482,677
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OSMAN, RICHARD A
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  UC93-300-4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 343-4341
;      TELEFAX:  (415) 343-4342
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2779 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-482-677-5


  Query Match             8.1%;  Score 18;  DB 3;  Length 2779;
  Best Local Similarity   100.0%;  Pred. No. 3.6;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      20 agtggggccgggacccgc 37
           ||||||||||||||||||
Db   1782 AGTGGGGCCGGGACCCGC 1765


RESULT   8
US-08-152-019A-41/c
; Sequence 41, Application US/08152019A
; Patent No. 5565331
;  GENERAL INFORMATION:
;    APPLICANT:  Tessier-Lavigne, Marc
;    APPLICANT:  Serafini, Tito
;    APPLICANT:  Kennedy, Timothy
;    APPLICANT:  Placzek, Marysia
;    APPLICANT:  Jessell, Thomas
;    APPLICANT:  Dodd, Jane
;    TITLE OF INVENTION:  NEURAL AXON OUTGROWTH MODULATORS
;    NUMBER OF SEQUENCES:  46
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/152,019A
;      FILING DATE:  12-NOV-1993
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Osman, Richard Aron
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  A-59012/RAO
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 781-1989
;      TELEFAX:  (415) 398-3249
;      TELEX:  910 277299 FHT UR
;  INFORMATION FOR SEQ ID NO:  41:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2783 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-152-019A-41


  Query Match             8.1%;  Score 18;  DB 1;  Length 2783;
  Best Local Similarity   100.0%;  Pred. No. 3.6;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      20 agtggggccgggacccgc 37
           ||||||||||||||||||
Db   1783 AGTGGGGCCGGGACCCGC 1766


RESULT   9
US-08-890-980-17
; Sequence 17, Application US/08890980
; Patent No. 5998141
;  GENERAL INFORMATION:
;    APPLICANT:  Acton, Susan L.
;    TITLE OF INVENTION:  SR-B1 NUCLEIC ACIDS AND USES THEREFOR
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FOLEY, HOAG & ELIOT LLP
;      STREET:  One Post Office Square
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109-2170
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/890,980
;      FILING DATE:  10-JUL-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Arnold, Beth E.
;      REGISTRATION NUMBER:  35,430
;      REFERENCE/DOCKET NUMBER:  MIA-005.01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-832-1000
;      TELEFAX:  617-832-7000
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  544 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-890-980-17


  Query Match             7.2%;  Score 16;  DB 2;  Length 544;
  Best Local Similarity   100.0%;  Pred. No. 37;
```

```
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    179 aaggggtaggagaaag 194
          ||||||||||||||||
Db    208 AAGGGGTAGGAGAAAG 223


RESULT  10
US-08-890-979-17
; Sequence 17, Application US/08890979
; Patent No. 6030778
;  GENERAL INFORMATION:
;    APPLICANT:  Acton, Susan L.
;    APPLICANT:  Ordovas, Jose M.
;    TITLE OF INVENTION:  DIAGNOSTIC ASSAYS AND KITS FOR BODY MASS
;    TITLE OF INVENTION:  DISORDERS
;    NUMBER OF SEQUENCES:  75
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FOLEY, HOAG & ELIOT LLP
;      STREET:  One Post Office Square
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109-2170
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/890,979
;      FILING DATE:  10-JUL-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Arnold, Beth E.
;      REGISTRATION NUMBER:  35,430
;      REFERENCE/DOCKET NUMBER:  MIA-005.02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-832-1000
;      TELEFAX:  617-832-7000
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  544 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-890-979-17


  Query Match             7.2%;  Score 16;  DB 3;  Length 544;
  Best Local Similarity   100.0%;  Pred. No. 37;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    179 aaggggtaggagaaag 194
          ||||||||||||||||
Db    208 AAGGGGTAGGAGAAAG 223


RESULT  11
US-09-032-894-17
; Sequence 17, Application US/09032894
; Patent No. 6130041
; GENERAL INFORMATION:
;  APPLICANT: Acton, Susan L.
;  TITLE OF INVENTION: SR-BI NUCLEIC ACIDS AND USES THEREFOR
;  FILE REFERENCE: MIA-005.03
;  CURRENT APPLICATION NUMBER: US/09/032,894
;  CURRENT FILING DATE: 1998-02-27
;  EARLIER APPLICATION NUMBER: 08/890,980
;  EARLIER FILING DATE: 1997-07-10
;  NUMBER OF SEQ ID NOS: 121
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
;   LENGTH: 544
;   TYPE: DNA
;   ORGANISM: Human
US-09-032-894-17


  Query Match             7.2%;  Score 16;  DB 3;  Length 544;
  Best Local Similarity   100.0%;  Pred. No. 37;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    179 aaggggtaggagaaag 194
          ||||||||||||||||
Db    208 aaggggtaggagaaag 223


RESULT  12
US-09-031-626-17
; Sequence 17, Application US/09031626
; Patent No. 6228581
; GENERAL INFORMATION:
;  APPLICANT: Acton, Susan L.
;  APPLICANT: Ordovas, Jose M.
;  TITLE OF INVENTION: DIAGNOSTIC ASSAYS AND KITS FOR BODY MASS AND
;  TITLE OF INVENTION: CARDIOVASCULAR DISORDERS
;  FILE REFERENCE: MIA-005.04
;  CURRENT APPLICATION NUMBER: US/09/031,626
;  CURRENT FILING DATE: 1998-02-27
;  EARLIER APPLICATION NUMBER: 08/890,979
;  EARLIER FILING DATE: 1997-07-10
;  NUMBER OF SEQ ID NOS: 121
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
;   LENGTH: 544
;   TYPE: DNA
;   ORGANISM: Human
US-09-031-626-17


  Query Match             7.2%;  Score 16;  DB 4;  Length 544;
  Best Local Similarity   100.0%;  Pred. No. 37;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    179 aaggggtaggagaaag 194
          ||||||||||||||||
Db    208 aaggggtaggagaaag 223


RESULT  13
US-08-890-980-5
; Sequence 5, Application US/08890980
; Patent No. 5998141
;  GENERAL INFORMATION:
;    APPLICANT:  Acton, Susan L.
;    TITLE OF INVENTION:  SR-B1 NUCLEIC ACIDS AND USES THEREFOR
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FOLEY, HOAG & ELIOT LLP
;      STREET:  One Post Office Square
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109-2170
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/890,980
;      FILING DATE:  10-JUL-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
```

```
;       NAME:  Arnold, Beth E.
;       REGISTRATION NUMBER:  35,430
;       REFERENCE/DOCKET NUMBER:  MIA-005.01
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  617-832-1000
;       TELEFAX:  617-832-7000
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1002 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-08-890-980-5


  Query Match             7.2%;  Score 16;  DB 2;  Length 1002;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    179 aaggggtaggagaaag 194
          ||||||||||||||||
Db    208 AAGGGGTAGGAGAAAG 223


RESULT  14
US-08-890-979-5
; Sequence 5, Application US/08890979
; Patent No. 6030778
;  GENERAL INFORMATION:
;     APPLICANT:  Acton, Susan L.
;     APPLICANT:  Ordovas, Jose M.
;     TITLE OF INVENTION:  DIAGNOSTIC ASSAYS AND KITS FOR BODY MASS
;     TITLE OF INVENTION:  DISORDERS
;     NUMBER OF SEQUENCES:  75
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  FOLEY, HOAG & ELIOT LLP
;       STREET:  One Post Office Square
;       CITY:  Boston
;       STATE:  MA
;       COUNTRY:  USA
;       ZIP:  02109-2170
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/890,979
;       FILING DATE:  10-JUL-1997
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Arnold, Beth E.
;       REGISTRATION NUMBER:  35,430
;       REFERENCE/DOCKET NUMBER:  MIA-005.02
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  617-832-1000
;       TELEFAX:  617-832-7000
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1002 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-08-890-979-5


  Query Match             7.2%;  Score 16;  DB 3;  Length 1002;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    179 aaggggtaggagaaag 194
          ||||||||||||||||
Db    208 AAGGGGTAGGAGAAAG 223


RESULT  15
US-09-032-894-5
; Sequence 5, Application US/09032894
; Patent No. 6130041
; GENERAL INFORMATION:
;  APPLICANT: Acton, Susan L.
;  TITLE OF INVENTION: SR-BI NUCLEIC ACIDS AND USES THEREFOR
;  FILE REFERENCE: MIA-005.03
;  CURRENT APPLICATION NUMBER: US/09/032,894
;  CURRENT FILING DATE: 1998-02-27
;  EARLIER APPLICATION NUMBER: 08/890,980
;  EARLIER FILING DATE: 1997-07-10
;  NUMBER OF SEQ ID NOS: 121
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 1002
;   TYPE: DNA
;   ORGANISM: Human
US-09-032-894-5


  Query Match             7.2%;  Score 16;  DB 3;  Length 1002;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    179 aaggggtaggagaaag 194
          ||||||||||||||||
Db    208 aaggggtaggagaaag 223


RESULT  16
US-09-032-894-95
; Sequence 95, Application US/09032894
; Patent No. 6130041
; GENERAL INFORMATION:
;  APPLICANT: Acton, Susan L.
;  TITLE OF INVENTION: SR-BI NUCLEIC ACIDS AND USES THEREFOR
;  FILE REFERENCE: MIA-005.03
;  CURRENT APPLICATION NUMBER: US/09/032,894
;  CURRENT FILING DATE: 1998-02-27
;  EARLIER APPLICATION NUMBER: 08/890,980
;  EARLIER FILING DATE: 1997-07-10
;  NUMBER OF SEQ ID NOS: 121
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 95
;   LENGTH: 1002
;   TYPE: DNA
;   ORGANISM: Human
US-09-032-894-95


  Query Match             7.2%;  Score 16;  DB 3;  Length 1002;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    179 aaggggtaggagaaag 194
          ||||||||||||||||
Db    208 aaggggtaggagaaag 223


RESULT  17
US-09-031-626-5
; Sequence 5, Application US/09031626
; Patent No. 6228581
; GENERAL INFORMATION:
;  APPLICANT: Acton, Susan L.
;  APPLICANT: Ordovas, Jose M.
;  TITLE OF INVENTION: DIAGNOSTIC ASSAYS AND KITS FOR BODY MASS AND
;  TITLE OF INVENTION: CARDIOVASCULAR DISORDERS
```

```
;  FILE REFERENCE: MIA-005.04
;  CURRENT APPLICATION NUMBER: US/09/031,626
;  CURRENT FILING DATE: 1998-02-27
;  EARLIER APPLICATION NUMBER: 08/890,979
;  EARLIER FILING DATE: 1997-07-10
;  NUMBER OF SEQ ID NOS: 121
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;    LENGTH: 1002
;    TYPE: DNA
;    ORGANISM: Human
US-09-031-626-5


  Query Match             7.2%;  Score 16;  DB 4;  Length 1002;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     179 aaggggtaggagaaag 194
           ||||||||||||||||
Db     208 aaggggtaggagaaag 223


RESULT  18
US-09-031-626-95
; Sequence 95, Application US/09031626
; Patent No. 6228581
; GENERAL INFORMATION:
;  APPLICANT: Acton, Susan L.
;  APPLICANT: Ordovas, Jose M.
;  TITLE OF INVENTION: DIAGNOSTIC ASSAYS AND KITS FOR BODY MASS AND
;  TITLE OF INVENTION: CARDIOVASCULAR DISORDERS
;  FILE REFERENCE: MIA-005.04
;  CURRENT APPLICATION NUMBER: US/09/031,626
;  CURRENT FILING DATE: 1998-02-27
;  EARLIER APPLICATION NUMBER: 08/890,979
;  EARLIER FILING DATE: 1997-07-10
;  NUMBER OF SEQ ID NOS: 121
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 95
;    LENGTH: 1002
;    TYPE: DNA
;    ORGANISM: Human
US-09-031-626-95


  Query Match             7.2%;  Score 16;  DB 4;  Length 1002;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     179 aaggggtaggagaaag 194
           ||||||||||||||||
Db     208 aaggggtaggagaaag 223


RESULT  19
US-09-105-537-21
; Sequence 21, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
;  APPLICANT: Sherman, D.H.
;  APPLICANT: Liu, H.
;  APPLICANT: Xue, Y.
;  APPLICANT: Zhao, L.
;  TITLE OF INVENTION: DNA encoding methymycin and pikromycin
;  FILE REFERENCE: 600.438US1
;  CURRENT APPLICATION NUMBER: US/09/105,537A
;  CURRENT FILING DATE: 1998-06-26
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 21
;    LENGTH: 1209
;    TYPE: DNA
;    ORGANISM: Streptomyces venezuelae
US-09-105-537-21


  Query Match             7.2%;  Score 16;  DB 4;  Length 1209;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      60 cccgggctgcggcagg 75
           ||||||||||||||||
Db    1120 cccgggctgcggcagg 1135


RESULT  20
US-08-318-905-21
; Sequence 21, Application US/08318905
; Patent No. 5641669
;  GENERAL INFORMATION:
;    APPLICANT:  Cousens, Lawrence S.
;    APPLICANT:  Eberhardt, Christine D.
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Le Trong, Hai
;    APPLICANT:  Tjoelker, Larry W.
;    APPLICANT:  Wilder, Cheryl L.
;    TITLE OF INVENTION:  Platelet-Activating Factor Acetyl
;    TITLE OF INVENTION:  Hydrolase
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gertein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/318,905
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/133,803
;      FILING DATE:  6-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5641669and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  32205
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3658
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1494 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  117..1436
US-08-318-905-21


  Query Match             7.2%;  Score 16;  DB 1;  Length 1494;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     206 caaaagtggaaaacag 221
           ||||||||||||||||
```

```
Db    700 CAAAAGTGGAAAACAG 715


RESULT  21
US-08-483-232-21
; Sequence 21, Application US/08483232
; Patent No. 5656431
;  GENERAL INFORMATION:
;    APPLICANT:  Cousens, Lawrence S.
;    APPLICANT:  Eberhardt, Christine D.
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Le Trong, Hai
;    APPLICANT:  Tjoelker, Larry W.
;    APPLICANT:  Wilder, Cheryl L.
;    TITLE OF INVENTION:  Platelet-Activating Factor
;    TITLE OF INVENTION:  Acetylhydrolase
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/483,232
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/318,905
;      FILING DATE:  06-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/133,803
;      FILING DATE:  06-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5656431and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/32689
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3658
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1494 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  117..1436
US-08-483-232-21


  Query Match             7.2%;  Score 16;  DB 1;  Length 1494;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     206 caaaagtggaaaacag 221
           ||||||||||||||||
Db     700 CAAAAGTGGAAAACAG 715


RESULT  22
US-08-483-140-21
; Sequence 21, Application US/08483140
; Patent No. 5698403
;  GENERAL INFORMATION:
;    APPLICANT:  ICOS Corporation
;    TITLE OF INVENTION:  Platelet-Activating Factor Acetyl
;    TITLE OF INVENTION:  Hydrolase
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gertein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/483,140
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/318,905
;      FILING DATE:  6-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/133,803
;      FILING DATE:  6-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5698403and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  32781
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3658
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1494 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  117..1436
US-08-483-140-21


  Query Match             7.2%;  Score 16;  DB 1;  Length 1494;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     206 caaaagtggaaaacag 221
           ||||||||||||||||
Db     700 CAAAAGTGGAAAACAG 715


RESULT  23
US-08-485-938A-21
; Sequence 21, Application US/08485938A
; Patent No. 5847088
;  GENERAL INFORMATION:
;    APPLICANT:  Cousens, Lawrence S.
;    APPLICANT:  Eberhardt, Christine D.
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Le Trong, Hai
;    APPLICANT:  Tjoelker, Larry W.
;    APPLICANT:  Wilder, Cheryl L.
;    TITLE OF INVENTION:  Platelet-Activating Factor
;    TITLE OF INVENTION:  Acetylhydrolase
;    NUMBER OF SEQUENCES:  36
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
```

```
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,938A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/318,905
;      FILING DATE:  06-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/133,803
;      FILING DATE:  06-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5847088and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/32792
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3658
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1494 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  117..1436
US-08-485-938A-21


  Query Match             7.2%;  Score 16;  DB 2;  Length 1494;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     206 caaaagtggaaaacag 221
           ||||||||||||||||
Db     700 CAAAAGTGGAAAACAG 715


RESULT  24
US-08-910-041-21
; Sequence 21, Application US/08910041
; Patent No. 5977308
;  GENERAL INFORMATION:
;    APPLICANT:  Cousens, Lawrence S.
;    APPLICANT:  Eberhardt, Christine D.
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Le Trong, Hai
;    APPLICANT:  Tjoelker, Larry W.
;    APPLICANT:  Wilder, Cheryl L.
;    TITLE OF INVENTION:  Platelet-Activating Factor
;    TITLE OF INVENTION:  Acetylhydrolase
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/910,041
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/483,232
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/318,905
;      FILING DATE:  06-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/133,803
;      FILING DATE:  06-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27866/34026
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3658
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1494 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  117..1436
US-08-910-041-21


  Query Match             7.2%;  Score 16;  DB 2;  Length 1494;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     206 caaaagtggaaaacag 221
           ||||||||||||||||
Db     700 CAAAAGTGGAAAACAG 715


RESULT  25
US-09-328-474-21
; Sequence 21, Application US/09328474
; Patent No. 6045794
;  GENERAL INFORMATION:
;    APPLICANT:  Cousens, Lawrence S.
;    APPLICANT:  Eberhardt, Christine D.
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Le Trong, Hai
;    APPLICANT:  Tjoelker, Larry W.
;    APPLICANT:  Wilder, Cheryl L.
;    TITLE OF INVENTION:  Platelet-Activating Factor
;    TITLE OF INVENTION:  Acetylhydrolase
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
```

```
;       APPLICATION NUMBER:  US/09/328,474
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/483,232
;       FILING DATE:  07-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/318,905
;       FILING DATE:  06-OCT-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/133,803
;       FILING DATE:  06-OCT-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Rin-Laures, Li-Hsien
;       REGISTRATION NUMBER:  33,547
;       REFERENCE/DOCKET NUMBER:  27866/34026
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (312) 474-6300
;       TELEFAX:  (312) 474-0448
;       TELEX:  25-3658
;   INFORMATION FOR SEQ ID NO:  21:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1494 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  117..1436
US-09-328-474-21


  Query Match             7.2%;  Score 16;  DB 3;  Length 1494;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     206 caaaagtggaaaacag 221
           ||||||||||||||||
Db     700 CAAAAGTGGAAAACAG 715


RESULT  26
US-09-100-546-21
; Sequence 21, Application US/09100546
; Patent No. 6099836
;  GENERAL INFORMATION:
;     APPLICANT:  Cousens, Lawrence S.
;     APPLICANT:  Eberhardt, Christine D.
;     APPLICANT:  Gray, Patrick W.
;     APPLICANT:  Le Trong, Hai
;     APPLICANT:  Tjoelker, Larry W.
;     APPLICANT:  Wilder, Cheryl L.
;     TITLE OF INVENTION:  Platelet-Activating Factor
;     TITLE OF INVENTION:  Acetylhydrolase
;     NUMBER OF SEQUENCES:  30
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;       STREET:  6300 Sears Tower, 233 South Wacker Drive
;       CITY:  Chicago
;       STATE:  Illinois
;       COUNTRY:  United States of America
;       ZIP:  60606-6402
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/100,546
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/010,715
;       FILING DATE:
;       APPLICATION NUMBER:  US 08/318,905
;       FILING DATE:  06-OCT-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/133,803
;       FILING DATE:  06-OCT-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  No. 6099836and, Greta E.
;       REGISTRATION NUMBER:  35,302
;       REFERENCE/DOCKET NUMBER:  27866/32793
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (312) 474-6300
;       TELEFAX:  (312) 474-0448
;       TELEX:  25-3658
;   INFORMATION FOR SEQ ID NO:  21:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1494 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  117..1436
US-09-100-546-21


  Query Match             7.2%;  Score 16;  DB 3;  Length 1494;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     206 caaaagtggaaaacag 221
           ||||||||||||||||
Db     700 CAAAAGTGGAAAACAG 715


RESULT  27
US-09-010-715-21
; Sequence 21, Application US/09010715
; Patent No. 6146625
;  GENERAL INFORMATION:
;     APPLICANT:  Cousens, Lawrence S.
;     APPLICANT:  Eberhardt, Christine D.
;     APPLICANT:  Gray, Patrick W.
;     APPLICANT:  Le Trong, Hai
;     APPLICANT:  Tjoelker, Larry W.
;     APPLICANT:  Wilder, Cheryl L.
;     TITLE OF INVENTION:  Platelet-Activating Factor
;     TITLE OF INVENTION:  Acetylhydrolase
;     NUMBER OF SEQUENCES:  30
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;       STREET:  6300 Sears Tower, 233 South Wacker Drive
;       CITY:  Chicago
;       STATE:  Illinois
;       COUNTRY:  United States of America
;       ZIP:  60606-6402
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/010,715
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/318,905
;       FILING DATE:  06-OCT-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/133,803
;       FILING DATE:  06-OCT-1993
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6146625and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/32793
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3658
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1494 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  117..1436
US-09-010-715-21


  Query Match             7.2%;  Score 16;  DB 3;  Length 1494;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     206 caaaagtggaaaacag 221
           ||||||||||||||||
Db     700 CAAAAGTGGAAAACAG 715


RESULT  28
US-09-577-758-21
; Sequence 21, Application US/09577758
; Patent No. 6203790
;  GENERAL INFORMATION:
;    APPLICANT:  Cousens, Lawrence S.
;    APPLICANT:  Eberhardt, Christine D.
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Le Trong, Hai
;    APPLICANT:  Tjoelker, Larry W.
;    APPLICANT:  Wilder, Cheryl L.
;    TITLE OF INVENTION:  Platelet-Activating Factor
;    TITLE OF INVENTION:  Acetylhydrolase
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/577,758
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/010,715
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/133,803
;      FILING DATE:  06-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6203790and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/32793
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3658
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1494 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  117..1436
US-09-577-758-21


  Query Match             7.2%;  Score 16;  DB 4;  Length 1494;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     206 caaaagtggaaaacag 221
           ||||||||||||||||
Db     700 CAAAAGTGGAAAACAG 715


RESULT  29
US-08-812-204-1
; Sequence 1, Application US/08812204
; Patent No. 5965790
;  GENERAL INFORMATION:
;    APPLICANT:  Acton, Susan L.
;    TITLE OF INVENTION:  SR-BI REGULATORY SEQUENCES AND
;    TITLE OF INVENTION:  THERAPEUTIC METHODS OF USE
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FOLEY, HOAG & ELIOT LLP
;      STREET:  One Post Office Square
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109-2170
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/812,204
;      FILING DATE:  06-MAR-1997
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Arnold, Beth E.
;      REGISTRATION NUMBER:  35,430
;      REFERENCE/DOCKET NUMBER:  MIA-014.01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-832-1000
;      TELEFAX:  617-832-7000
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1613 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-812-204-1


  Query Match             7.2%;  Score 16;  DB 2;  Length 1613;
  Best Local Similarity   100.0%;  Pred. No. 35;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     179 aaggggtaggagaaag 194
           ||||||||||||||||
Db    1258 AAGGGGTAGGAGAAAG 1273
```

```
RESULT  30
PCT-US94-09752-2
; Sequence 2, Application PC/TUS9409752
;  GENERAL INFORMATION:
;     APPLICANT:  David S. Strayer and Avinash Chander
;     TITLE OF INVENTION:  Compositions and Methods for
;     TITLE OF INVENTION:  Targeting Cells and Modulating Pulmonary Surfactant Secretion
;     NUMBER OF SEQUENCES:  4
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Jane Massey Licata, Esq.
;       STREET:  210 Lake Drive East, Suite 201
;       CITY:  Cherry Hill
;       STATE:  NJ
;       COUNTRY:  USA
;       ZIP:  08002
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb
;       MEDIUM TYPE:  STORAGE
;       COMPUTER:  IBM PS/2
;       OPERATING SYSTEM:  PC-DOS
;       SOFTWARE:  WORDPERFECT 5.1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US94/09752
;       FILING DATE:  Herewith
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/176,218
;       FILING DATE:  December 30, 1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/114,951
;       FILING DATE:  August 31, 1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Jane Massey Licata
;       REGISTRATION NUMBER:  32,257
;       REFERENCE/DOCKET NUMBER:  JEFF-0042
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (609) 779-2400
;       TELEFAX:  (609) 779-8488
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1881
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single stranded
;       TOPOLOGY:  linear
PCT-US94-09752-2


  Query Match             7.2%;  Score 16;  DB 5;  Length 1881;
  Best Local Similarity   100.0%;  Pred. No. 35;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     172 gcttgtcaaggggtag 187
           ||||||||||||||||
Db     852 GCTTGTCAAGGGGTAG 867


RESULT  31
US-07-932-454A-2
; Sequence 2, Application US/07932454A
; Patent No. 5262318
;  GENERAL INFORMATION:
;     APPLICANT:  GUTHRIE, ELLEN P.
;     TITLE OF INVENTION:  ISOLATED DNA ENCODING THE SphI
;     TITLE OF INVENTION:  RESTRICTION ENDONUCLEASE AND RELATED METHODS FOR PRODUCING
;     TITLE OF INVENTION:  THE SAME
;     NUMBER OF SEQUENCES:  5
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
;       ADDRESSEE:  CUSHMAN
;       STREET:  130 WATER STREET
;       CITY:  BOSTON
;       STATE:  MASSACHUSETTS
;       COUNTRY:  US
;       ZIP:  02109
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/932,454A
;       FILING DATE:  19920820
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  WILLIAMS, GREGORY D.
;       REGISTRATION NUMBER:  30901
;       REFERENCE/DOCKET NUMBER:  42078
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 523-3400
;       TELEFAX:  (617) 5523-6440
;       TELEX:  200291 STRE UR
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2692 base pairs
;       TYPE:  NUCLEIC ACID
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  703..1653
;       OTHER INFORMATION:  /note= "METHYLASE GENE STARTS AT
;       OTHER INFORMATION:  POSITION 703/ENDS AT 1653.  RESTRICTION
;       OTHER INFORMATION:  ENDONUCLEASE STARTS AT POSITION 1703/ENDS AT 2410"
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  1703..2410
US-07-932-454A-2


  Query Match             7.2%;  Score 16;  DB 1;  Length 2692;
  Best Local Similarity   100.0%;  Pred. No. 35;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      21 gtggggccgggacccg 36
           ||||||||||||||||
Db     462 GTGGGGCCGGGACCCG 477


RESULT  32
US-08-980-241-8/c
; Sequence 8, Application US/08980241D
; Patent No. 6319708
; GENERAL INFORMATION:
;  APPLICANT:  Chalfie, Martin
;  APPLICANT:  Taub, James J.
;  TITLE OF INVENTION:  A METHOD FOR INCREASING LIFE-SPAN
;  FILE REFERENCE:  0575/51778/JPW/JSG
;  CURRENT APPLICATION NUMBER: US/08/980,241D
;  CURRENT FILING DATE:  1997-11-28
;  NUMBER OF SEQ ID NOS:  8
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 8
;   LENGTH: 6840
;   TYPE: DNA
;   ORGANISM: Nematodes
;   FEATURE:
;   NAME/KEY: N_region
;   LOCATION: (2138)
;   OTHER INFORMATION: N= g, a, c or t(u)
;   NAME/KEY: N_region
;   LOCATION: (3054)
;   OTHER INFORMATION: N=  g, a, c  or t(u)
;   NAME/KEY: N_region
;   LOCATION: (3060)
;   OTHER INFORMATION: N= g, a, c  or t(u)
```

```
;   NAME/KEY: N_region
;   LOCATION: (3070)
;   OTHER INFORMATION: N= g, a, c or t(u)
;   NAME/KEY: N_region
;   LOCATION: (3905)
;   OTHER INFORMATION: N= g, a, c or t(u)
;   NAME/KEY: N_region
;   LOCATION: (3913)
;   OTHER INFORMATION: N= g, a, c or t(u)
;   NAME/KEY: N_region
;   LOCATION: (3917)
;   OTHER INFORMATION: N= g, a, c or t(u)
;   NAME/KEY: N_region
;   LOCATION: (4045)
;   OTHER INFORMATION: N= g, a, c or t(u)
;   NAME/KEY: N_region
;   LOCATION: (4412)..(4413)
;   OTHER INFORMATION: N= g, a,  c or t(u)
;   NAME/KEY: N_region
;   LOCATION: (4416)
;   OTHER INFORMATION: N= g, a,  c or t(u)
;   NAME/KEY: N_region
;   LOCATION: (4735)
;   OTHER INFORMATION: N= g, a, c  or t(u)
;   NAME/KEY: N_region
;   LOCATION: (4876)
;   OTHER INFORMATION: N= g, a, c or t(u)
US-08-980-241-8


  Query Match             7.2%;  Score 16;  DB 4;  Length 6840;
  Best Local Similarity   100.0%;  Pred. No. 34;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     187 ggagaaagagacgcaa 202
           ||||||||||||||||
Db    5890 GGAGAAAGAGACGCAA 5875


RESULT  33
US-09-105-537-3
; Sequence 3, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
;  APPLICANT: Sherman, D.H.
;  APPLICANT: Liu, H.
;  APPLICANT: Xue, Y.
;  APPLICANT: Zhao, L.
;  TITLE OF INVENTION: DNA encoding methymycin and pikromycin
;  FILE REFERENCE: 600.438US1
;  CURRENT APPLICATION NUMBER: US/09/105,537A
;  CURRENT FILING DATE: 1998-06-26
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 13613
;   TYPE: DNA
;   ORGANISM: Streptomyces venezuelae
US-09-105-537-3


  Query Match             7.2%;  Score 16;  DB 4;  Length 13613;
  Best Local Similarity   100.0%;  Pred. No. 33;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      60 cccgggctgcggcagg 75
           ||||||||||||||||
Db    1925 cccgggctgcggcagg 1940


RESULT  34
US-08-922-635-21
; Sequence 21, Application US/08922635A
; Patent No. 6033871
; GENERAL INFORMATION:
;  APPLICANT: PILETZ, John E.
;  APPLICANT: IVANOV, Tina R.
;  TITLE OF INVENTION: DNA MOLECULES ENCODING IMIDALINE RECEPTIVE POLYPEPTIDES
;  TITLE OF INVENTION: AND POLYPEPTIDES ENCODED THEREBY
;  FILE REFERENCE: Corrected Sequence Listing
; Patent No. 6033871
;  CURRENT APPLICATION NUMBER: US/08/922,635A
;  CURRENT FILING DATE: 1997-09-03
;  EARLIER APPLICATION NUMBER: 08/650,766
;  EARLIER FILING DATE: 1996-05-20
;  EARLIER APPLICATION NUMBER: 60/012,600
;  EARLIER FILING DATE: 1996-03-01
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
;   LENGTH: 15202
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-08-922-635-21


  Query Match             7.2%;  Score 16;  DB 3;  Length 15202;
  Best Local Similarity   100.0%;  Pred. No. 33;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      53 cttctctcccgggctg 68
           ||||||||||||||||
Db    1370 cttctctcccgggctg 1385


RESULT  35
US-09-147-236-1/c
; Sequence 1, Application US/09147236A
; Patent No. 6316251
; GENERAL INFORMATION:
;  APPLICANT: TONOUCHI, Naoto
;  APPLICANT: TSUCHIDA, Takayasu
;  APPLICANT: YOSHINAGA, Fumihiro
;  APPLICANT: TAHARA, Naoki
;  APPLICANT: HAYASHI, Takahisa
;  TITLE OF INVENTION: NOVEL GENE, GROUP OF GENES, AND NOVEL BETA-GLUCOSIDASE
;  FILE REFERENCE: 6537-011-0PCT
;  CURRENT APPLICATION NUMBER: US/09/147,236A
;  CURRENT FILING DATE: 1999-04-08
;  EARLIER APPLICATION NUMBER: PCT/JP97/03633
;  EARLIER FILING DATE: 1997-10-09
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 16836
;   TYPE: DNA
;   ORGANISM: Acetobacter xylinum
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (869)..(1891)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (3101)..(5368)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (5373)..(7778)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (7784)..(11761)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (11764)..(12231)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (12448)..(14652)
;   FEATURE:
```

```
;   OTHER INFORMATION: n at positions 15741 and 15767 may be a, g, t, or
;   OTHER INFORMATION: c
US-09-147-236-1


  Query Match             7.2%;  Score 16;  DB 4;  Length 16836;
  Best Local Similarity   100.0%;  Pred. No. 33;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      62 cgggctgcggcagggc 77
           ||||||||||||||||
Db   14919 CGGGCTGCGGCAGGGC 14904


RESULT  36
US-09-147-236-10/c
; Sequence 10, Application US/09147236A
; Patent No. 6316251
; GENERAL INFORMATION:
;  APPLICANT: TONOUCHI, Naoto
;  APPLICANT: TSUCHIDA, Takayasu
;  APPLICANT: YOSHINAGA, Fumihiro
;  APPLICANT: TAHARA, Naoki
;  APPLICANT: HAYASHI, Takahisa
;  TITLE OF INVENTION: NOVEL GENE, GROUP OF GENES, AND NOVEL BETA-GLUCOSIDASE
;  FILE REFERENCE: 6537-011-0PCT
;  CURRENT APPLICATION NUMBER: US/09/147,236A
;  CURRENT FILING DATE: 1999-04-08
;  EARLIER APPLICATION NUMBER: PCT/JP97/03633
;  EARLIER FILING DATE: 1997-10-09
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 16836
;   TYPE: DNA
;   ORGANISM: Acetobacter xylinum
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1891)..(2922)
;   FEATURE:
;   OTHER INFORMATION: Nucleotide sequence is the same as SEQ ID NO:1
;   FEATURE:
;   OTHER INFORMATION: n at positions 15741 and 15767 may be a, g, c, or
;   OTHER INFORMATION: t
US-09-147-236-10


  Query Match             7.2%;  Score 16;  DB 4;  Length 16836;
  Best Local Similarity   100.0%;  Pred. No. 33;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      62 cgggctgcggcagggc 77
           ||||||||||||||||
Db   14919 CGGGCTGCGGCAGGGC 14904


RESULT  37
US-09-320-878-19
; Sequence 19, Application US/09320878A
; Patent No. 6117659
; GENERAL INFORMATION:
;  APPLICANT: ASHLEY, Gary
;  APPLICANT: BETLACH, Melanie C.
;  APPLICANT: BETLACH, Mary C.
;  APPLICANT: McDANIEL, Robert
;  APPLICANT: TANG, Li
;  TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
;  FILE REFERENCE: 300622002120
;  CURRENT APPLICATION NUMBER: US/09/320,878A
;  CURRENT FILING DATE: 1999-05-27
;  EARLIER APPLICATION NUMBER: CIP OF 09/141,908
;  EARLIER FILING DATE: 1998-08-28
;  EARLIER APPLICATION NUMBER: CIP OF 09/073,538
;  EARLIER FILING DATE: 1998-05-06
;  EARLIER APPLICATION NUMBER: CIP OF 08/846,247
;  EARLIER FILING DATE: 1997-04-30
;  EARLIER APPLICATION NUMBER: 60/119,139
;  EARLIER FILING DATE: 1999-02-08
;  EARLIER APPLICATION NUMBER: 60/100,880
;  EARLIER FILING DATE: 1998-09-22
;  EARLIER APPLICATION NUMBER: 60/087,080
;  EARLIER FILING DATE: 1998-05-28
;  NUMBER OF SEQ ID NOS: 34
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
;   LENGTH: 38506
;   TYPE: DNA
;   ORGANISM: Streptomyces venezuelae
US-09-320-878-19


  Query Match             7.2%;  Score 16;  DB 3;  Length 38506;
  Best Local Similarity   100.0%;  Pred. No. 32;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      60 cccgggctgcggcagg 75
           ||||||||||||||||
Db   35922 cccgggctgcggcagg 35937


RESULT  38
US-09-049-714-1/c
; Sequence 1, Application US/09049714
; Patent No. 6309827
;  GENERAL INFORMATION:
;    APPLICANT:  Goldstein, Andrew S.
;    APPLICANT:  Bestwick, Richard K.
;    TITLE OF INVENTION:  Simultaneous Collection of DNA and
;    TITLE OF INVENTION:  No. 6309827-Nucleic Acid Analytes
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, 8th Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/049,714
;      FILING DATE:  27-MAR-1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/042,124
;      FILING DATE:  28-MAR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weber, Kenneth A.
;      REGISTRATION NUMBER:  31,677
;      REFERENCE/DOCKET NUMBER:  017197-002410US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-09-049-714-1
```

```
  Query Match             6.8%;  Score 15;  DB 4;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    193 agagacgcaaacaca 207
          |||||||||||||||
Db     18 AGAGACGCAAACACA 4


RESULT  39
US-08-928-465-3/c
; Sequence 3, Application US/08928465
; Patent No. 6204024
;  GENERAL INFORMATION:
;    APPLICANT:  Romano, Joseph
;    APPLICANT:  Lee, Eun Mi
;    TITLE OF INVENTION:  CCR5 RNA Transcription Based
;    TITLE OF INVENTION:  Amplification Assay
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 6204024el Patent Department
;      STREET:  1300 Piccard Drive
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/928,465
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gormley, Mary E.
;      REGISTRATION NUMBER:  34,409
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-948-7400
;      TELEFAX:  301-948-9751
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "DNA Oligonucleotide"
;    HYPOTHETICAL:  NO
US-08-928-465-3


  Query Match             6.8%;  Score 15;  DB 4;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    193 agagacgcaaacaca 207
          |||||||||||||||
Db     18 AGAGACGCAAACACA 4


RESULT  40
US-09-481-288-2/c
; Sequence 2, Application US/09481288
; Patent No. 6235504
; GENERAL INFORMATION:
;  APPLICANT:  Zhang, Linqi
;  APPLICANT:  Lewin, Sharon R
;  APPLICANT:  Kostrikis, Leondios
;  APPLICANT:  Ho, David D
;  TITLE OF INVENTION:  METHOD FOR IDENTIFYING GENOMIC EQUIVALENT MARKERS AND
;  TITLE OF INVENTION:  THEIR USE IN QUANTITATING CELLS AND POLYNUCLEOTIDE
;  TITLE OF INVENTION:  SEQUENCES THEREIN
;  FILE REFERENCE:  2378-1-001N
;  CURRENT APPLICATION NUMBER: US/09/481,288
;  CURRENT FILING DATE:  2000-01-11
;  NUMBER OF SEQ ID NOS:  29
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 24
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: PRIMER
US-09-481-288-2


  Query Match             6.8%;  Score 15;  DB 4;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    193 agagacgcaaacaca 207
          |||||||||||||||
Db     17 AGAGACGCAAACACA 3


RESULT  41
US-08-850-049-108/c
; Sequence 108, Application US/08850049
; Patent No. 5965726
;  GENERAL INFORMATION:
;    APPLICANT:
;    APPLICANT:
;    APPLICANT:
;    TITLE OF INVENTION:  METHOD OF ELIMINATING
;    TITLE OF INVENTION:  INHIBITORY/INSTABILITY REGIONS OF mRNA
;    NUMBER OF SEQUENCES:  130
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/850,049
;      FILING DATE:  02-MAY-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/050,478
;      FILING DATE:  26-OCT-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/02908
;      FILING DATE:  29-MAR-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/858,747
;      FILING DATE:  27-MAR-1992
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MORRY, MARY J.
;      REGISTRATION NUMBER:  34,398
;      REFERENCE/DOCKET NUMBER:  2026-4006US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)758-4800
;      TELEFAX:  (212)751-6849
;  INFORMATION FOR SEQ ID NO:  108:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  43 BASE PAIRS
```

```
;       TYPE:  NUCLEIC ACID
;       STRANDEDNESS:  SINGLE
;       TOPOLOGY:  LINEAR
US-08-850-049-108


  Query Match             6.8%;  Score 15;  DB 2;  Length 43;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      67 tgcggcagggcaggg 81
           ||||||||||||||||
Db      23 TGCGGCAGGGCAGGG 9


RESULT  42
US-08-050-478-108/c
; Sequence 108, Application US/08050478
; Patent No. 5972596
;  GENERAL INFORMATION:
;    APPLICANT:
;    APPLICANT:
;    APPLICANT:
;    TITLE OF INVENTION:  METHOD OF ELIMINATING
;    TITLE OF INVENTION:  INHIBITORY/INSTABILITY REGIONS OF mRNA
;    NUMBER OF SEQUENCES:  130
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/050,478
;      FILING DATE:  26-OCT-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/02908
;      FILING DATE:  29-MAR-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/858,747
;      FILING DATE:  27-MAR-1992
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MORRY, MARY J.
;      REGISTRATION NUMBER:  34,398
;      REFERENCE/DOCKET NUMBER:  2026-4006US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)758-4800
;      TELEFAX:  (212)751-6849
;  INFORMATION FOR SEQ ID NO:  108:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  43 BASE PAIRS
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  LINEAR
US-08-050-478-108


  Query Match             6.8%;  Score 15;  DB 2;  Length 43;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      67 tgcggcagggcaggg 81
           ||||||||||||||||
Db      23 TGCGGCAGGGCAGGG 9


RESULT  43
US-09-414-117-108/c
; Sequence 108, Application US/09414117
; Patent No. 6291664
;  GENERAL INFORMATION:
;    APPLICANT:
;    APPLICANT:
;    APPLICANT:
;    TITLE OF INVENTION:  METHOD OF ELIMINATING
;    TITLE OF INVENTION:  INHIBITORY/INSTABILITY REGIONS OF mRNA
;    NUMBER OF SEQUENCES:  130
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/414,117
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/850,049
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/02908
;      FILING DATE:  29-MAR-1993
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/858,747
;      FILING DATE:  27-MAR-1992
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MORRY, MARY J.
;      REGISTRATION NUMBER:  34,398
;      REFERENCE/DOCKET NUMBER:  2026-4006US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)758-4800
;      TELEFAX:  (212)751-6849
;  INFORMATION FOR SEQ ID NO:  108:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  43 BASE PAIRS
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  LINEAR
US-09-414-117-108


  Query Match             6.8%;  Score 15;  DB 4;  Length 43;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      67 tgcggcagggcaggg 81
           ||||||||||||||||
Db      23 TGCGGCAGGGCAGGG 9


RESULT  44
US-09-046-247-24
; Sequence 24, Application US/09046247
; Patent No. 6124449
;  GENERAL INFORMATION:
;    APPLICANT:  NIKOS PAGRATIS
;    APPLICANT:  LARRY GOLD
```

```
;    TITLE OF INVENTION:  HIGH AFFINITY TGF? NUCLEIC
;    NUMBER OF SEQUENCES:  143
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Swanson and Bratschun, L.L.C.
;      STREET:  8400 East Prentice Avenue, Suite #200
;      CITY:  Denver
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 1.44 Mb storage
;      COMPUTER:  IBM compatible
;      OPERATING SYSTEM:  MS DOS
;      SOFTWARE:  Word 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/046,247
;      FILING DATE:  23-MARCH-1998
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/458,424
;      FILING DATE:  2-JUNE-1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/714,131
;      FILING DATE:  10-JUNE-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/536,428
;      FILING DATE:  11-JUNE-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/964,624
;      FILING DATE:  21-OCTOBER-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/117,991
;      FILING DATE:  8-SEPTEMBER-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/931,473
;      FILING DATE:  17-AUGUST-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barry Swanson
;      REGISTRATION NUMBER:  33,215
;      REFERENCE/DOCKET NUMBER:  NEX 34.2/CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 793-3333
;      TELEFAX:  (303) 793-3433
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  51 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  RNA
;    FEATURE:
;      OTHER INFORMATION:  All pyrimidines are 2'-F modified
US-09-046-247-24


  Query Match             6.8%;  Score 15;  DB 3;  Length 51;
  Best Local Similarity   46.7%;  Pred. No. 1.3e+02;
  Matches    7;  Conservative    8;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     146 ttctgttttctcccc 160
           ::|:|::::|:||||
Db      22 UUCUGUUUUCUCCCC 36


RESULT  45
US-09-481-288-1/c
; Sequence 1, Application US/09481288
; Patent No. 6235504
; GENERAL INFORMATION:
;  APPLICANT:  Zhang, Linqi
;  APPLICANT:  Lewin, Sharon R
;  APPLICANT:  Kostrikis, Leondios
;  APPLICANT:  Ho, David D
;  TITLE OF INVENTION:  METHOD FOR IDENTIFYING GENOMIC EQUIVALENT MARKERS AND
;  TITLE OF INVENTION:  THEIR USE IN QUANTITATING CELLS AND POLYNUCLEOTIDE
;  TITLE OF INVENTION:  SEQUENCES THEREIN
;  FILE REFERENCE:  2378-1-001N
;  CURRENT APPLICATION NUMBER: US/09/481,288
;  CURRENT FILING DATE:  2000-01-11
;  NUMBER OF SEQ ID NOS:  29
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 239
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-481-288-1


  Query Match             6.8%;  Score 15;  DB 4;  Length 239;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     193 agagacgcaaacaca 207
           |||||||||||||||
Db      17 AGAGACGCAAACACA 3


Search completed: September 20, 2002, 06:15:34
Job time: 11153 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 04:07:24 ; Search time 3900.56 Seconds
(without alignments)
764.718 Million cell updates/sec

Title: US-09-846-456-4
Perfect score: 221
Sequence: 1 gtaattgcgagcgagagtga.....aacacaaaagtggaaaacag 221

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	219	99.1	763	9	AU121731	AU121731 AU121731
2	217	98.2	736	9	AU135588	AU135588 AU135588
3	154	69.7	292	10	Z44377	Z44377 HSC1ZB081 n
4	98	44.3	998	10	BG678861	BG678861 602624760
5	30	13.6	535	10	BG384217	BG384217 303216 MA
c 6	21	9.5	679	12	AG037352	AG037352 Pan trogl
7	21	9.5	710	12	BH502182	BH502182 BOGHN04TR
8	20	9.0	265	9	BB862540	BB862540 BB862540
9	20	9.0	389	10	BG629097	BG629097 cC-esflcL
10	20	9.0	426	10	BG629091	BG629091 cC-esflcL
c 11	20	9.0	777	10	BG123629	BG123629 EST469275
c 12	19	8.6	87	9	AU076672	AU076672 AU076672
c 13	19	8.6	300	9	AU099072	AU099072 AU099072
c 14	19	8.6	389	10	BI772330	BI772330 603056038
15	19	8.6	484	9	AW642780	AW642780 cm22e02.w
16	19	8.6	486	10	T78163	T78163 yd79b06.r1
c 17	19	8.6	530	10	BG702824	BG702824 602684727
c 18	19	8.6	554	9	AL597062	AL597062 DKFZp313C
c 19	19	8.6	597	10	BI548754	BI548754 603189147
c 20	19	8.6	609	10	BI523553	BI523553 603175746
c 21	19	8.6	616	10	BG714492	BG714492 602670936
c 22	19	8.6	656	10	BG776309	BG776309 602663377
c 23	19	8.6	664	10	BM311005	BM311005 ig59e10.y
c 24	19	8.6	674	10	BI561894	BI561894 603255765
c 25	19	8.6	686	10	BI545327	BI545327 603187458
26	19	8.6	688	10	BI549570	BI549570 603192272
c 27	19	8.6	694	10	BJ044199	BJ044199 BJ044199
c 28	19	8.6	703	10	BJ052576	BJ052576 BJ052576
c 29	19	8.6	704	9	AL554654	AL554654 AL554654
c 30	19	8.6	706	10	BG431399	BG431399 602500027
c 31	19	8.6	725	9	AL600963	AL600963 DKFZp313B
c 32	19	8.6	730	10	BE973918	BE973918 601680234
c 33	19	8.6	731	10	BI824845	BI824845 603033758
c 34	19	8.6	748	10	BG564438	BG564438 602584384
c 35	19	8.6	751	9	AL601135	AL601135 DKFZp313O
c 36	19	8.6	756	10	BI088379	BI088379 602851164
c 37	19	8.6	757	10	BG706900	BG706900 602672070
c 38	19	8.6	781	9	AU143746	AU143746 AU143746
c 39	19	8.6	784	10	BI458234	BI458234 603199174
c 40	19	8.6	795	10	BG705539	BG705539 602668716
c 41	19	8.6	803	10	BI548062	BI548062 603196572
c 42	19	8.6	813	9	AL542978	AL542978 AL542978
c 43	19	8.6	824	10	BG402101	BG402101 602465634
c 44	19	8.6	848	10	BI600654	BI600654 603247539
c 45	19	8.6	879	10	BI764049	BI764049 603043255

ALIGNMENTS

RESULT 1
AU121731
LOCUS AU121731 763 bp mRNA linear EST 19-OCT-2000
DEFINITION AU121731 MAMMA1 Homo sapiens cDNA clone MAMMA1000851 5', mRNA sequence.
ACCESSION AU121731
VERSION AU121731.1 GI:10936966
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 763)
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.
FEATURES Location/Qualifiers
source 1. .763
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MAMMA1000851"
/clone_lib="MAMMA1"
/tissue_type="mammary gland"
/note="Vector: pME18SFL3"
BASE COUNT 137 a 205 c 260 g 158 t 3 others
ORIGIN

```
Query Match           99.1%;  Score 219;  DB 9;  Length 763;
Best Local Similarity  100.0%;  Pred. No. 6.4e-101;
Matches  219;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 aattgcgagcgagagtgagtggggccgggacccgcagagccgagccgaccctctctccc 62
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 AATTGCGAGCGAGAGTGAGTGGGGCCGGGACCCGCAGAGCCGAGCCGACCCTTCTCTCCC 60

Qy      63 gggctgcggcagggcagggcggggagctccgcgcaccaacagagccggttctcagggcgc 122
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GGGCTGCGGCAGGGCAGGGCGGGGAGCTCCGCGCACCAACAGAGCCGGTTCTCAGGGCGC 120

Qy     123 tttgctccttgtttttttccccggttctgttttctcccccttctccggaaggcttgtcaagg 182
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TTTGCTCCTTGTTTTTTCCCCGGTTCTGTTTTCTCCCCTTCTCCGGAAGGCTTGTCAAGG 180

Qy     183 ggtaggagaaagagacgcaaacacaaaagtggaaaacag 221
           |||||||||||||||||||||||||||||||||||||||
Db     181 GGTAGGAGAAAGAGACGCAAACACAAAAGTGGAAAACAG 219
```

```
RESULT   2
AU135588
LOCUS       AU135588                 736 bp    mRNA    linear   EST 24-OCT-2000
DEFINITION  AU135588 PLACE1 Homo sapiens cDNA clone PLACE1002437 5', mRNA
            sequence.
ACCESSION   AU135588
VERSION     AU135588.1  GI:10996127
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 736)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .736
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="PLACE1002437"
                     /clone_lib="PLACE1"
                     /tissue_type="placenta"
                     /note="Vector: pME18SFL3"
BASE COUNT       163 a    199 c    199 g    170 t      5 others
ORIGIN
```

```
Query Match           98.2%;  Score 217;  DB 9;  Length 736;
Best Local Similarity  100.0%;  Pred. No. 6.7e-100;
Matches  217;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 ttgcgagcgagagtgagtggggccgggacccgcagagccgagccgaccctctctcccgg 64
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5 TTGCGAGCGAGAGTGAGTGGGGCCGGGACCCGCAGAGCCGAGCCGACCCTTCTCTCCCGG 64

Qy      65 gctgcggcagggcagggcggggagctccgcgcaccaacagagccggttctcagggcgctt 124
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      65 GCTGCGGCAGGGCAGGGCGGGGAGCTCCGCGCACCAACAGAGCCGGTTCTCAGGGCGCTT 124

Qy     125 tgctccttgttttttccccggttctgttttctcccccttctccggaaggcttgtcaagggg 184
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125 TGCTCCTTGTTTTTTCCCCGGTTCTGTTTTCTCCCCTTCTCCGGAAGGCTTGTCAAGGGG 184

Qy     185 taggagaaagagacgcaaacacaaaagtggaaaacag 221
           |||||||||||||||||||||||||||||||||||||
Db     185 TAGGAGAAAGAGACGCAAACACAAAAGTGGAAAACAG 221
```

```
RESULT   3
Z44377
LOCUS       Z44377                   292 bp    mRNA    linear   EST 14-NOV-1994
DEFINITION  HSC1ZB081 normalized infant brain cDNA Homo sapiens cDNA clone
            c-1zb08, mRNA sequence.
ACCESSION   Z44377
VERSION     Z44377.1  GI:573506
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 292)
  AUTHORS   Auffray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C., Devignes
            ,M.D., Duprat,S., Houlgatte,R., Jumeau,M.N., Lamy,B., Lorenzo,F.,
            Mitchell,H., Mariage-Samson,R., Pietu,G., Pouliot,Y.,
            Sebastiani-Kabaktchis,C. and Tessier,A.
  TITLE     IMAGE: molecular integration of the analysis of the human genome
            and its expression
  JOURNAL   C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
  MEDLINE   95277534
COMMENT     Contact: Genethon
            Genexpress-Genethon
            Genethon Centre de recherche sur le Genome Humain
            1,rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
            Tel: 33169472800
            Fax: 33160778698
            Email: genexpress@genethon.fr
            Single read.
            Genexpress_library_idt: C; Genexpress_sequence_idt: y1c-1zb08
            Seq primer: (-21)M13_universal.
FEATURES             Location/Qualifiers
     source          1. .292
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="c-1zb08"
                     /clone_lib="normalized infant brain cDNA"
                     /sex="Female"
                     /tissue_type="total brain"
                     /dev_stage="3 months old"
                     /note="Organ: brain; Vector: lafmid BA; Site_1: HindIII;
                     Site_2: NotI; sex=Female; dev_stage=3 months old;
                     isolate=muscular atrophy patient;  tissue_type=total brain
                     ;  total mRNA was oligo-(dT) primed and directionally
                     cloned 5' -> 3' into the  HindIII -> NotI sites of the
                     lafmid BA vector. Clone library from B.Soares,  Psychiatry
                     Dept. Columbia University, USA.  Normalization_method:
                     Bento Soares, P.N.A.S in press"
BASE COUNT        50 a     87 c     96 g     56 t      3 others
ORIGIN
```

```
Query Match           69.7%;  Score 154;  DB 10;  Length 292;
Best Local Similarity  99.5%;  Pred. No. 8.9e-68;
Matches  204;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      17 gtgagtggggccgggacccgcagagccgagccgaccctctctctcccgggctgcggcaggg 76
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GTGAGTGGGGCCGGGACCCGCAGAGCCGAGCCGACCCTTCTCTCCCGGGCTGCGGCAGGG 60

Qy      77 cagggcggggagctccgcgcaccaacagagccggttctcagggcgctttgctccttgttt 136
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     61 CAGGGCGGGGAGCTCCGCGCACCAACAGAGCCGGTTCTCAGGGCGCTTTGCTCCTTGTTT 120

Qy    137 tttccccggttctgttttctccccttctccggaaggcttgtcaaggggtaggagaaagag 196
          ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db    121 TTTCCCCGGTTCTGTTTTCTCCCCTTNTCCGGAAGGCTTGTCAAGGGGTAGGAGAAAGAG 180

Qy    197 acgcaaacacaaaagtggaaaacag 221
          |||||||||||||||||||||||||
Db    181 ACGCAAACACAAAAGTGGAAAACAG 205


RESULT   4
BG678861
LOCUS       BG678861                998 bp    mRNA    linear   EST 01-MAY-2001
DEFINITION  602624760F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4749735 5',
            mRNA sequence.
ACCESSION   BG678861
VERSION     BG678861.1  GI:13910258
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 998)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Cleaver, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10603  row: g  column: 16
            High quality sequence stop: 860.
FEATURES             Location/Qualifiers
     source          1. .998
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4749735"
                     /clone_lib="NCI_CGAP_Skn4"
                     /tissue_type="squamous cell carcinoma"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.5kb.  Library  constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      285 a    233 c    244 g    236 t
ORIGIN


  Query Match           44.3%;  Score 98;  DB 10;  Length 998;
  Best Local Similarity  100.0%;  Pred. No. 2.7e-39;
  Matches   98;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    124 ttgctccttgtttttttccccggttctgttttctcccccttctccggaaggcttgtcaaggg 183
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105 TTGCTCCTTGTTTTTTCCCCGGTTCTGTTTTCTCCCCTTCTCCGGAAGGCTTGTCAAGGG 164

Qy    184 gtaggagaaagagacgcaaacacaaaagtggaaaacag 221
          |||||||||||||||||||||||||||||||||||||
Db    165 GTAGGAGAAAGAGACGCAAACACAAAAGTGGAAAACAG 202


RESULT   5
BG384217
LOCUS       BG384217                535 bp    mRNA    linear   EST 12-MAR-2001
DEFINITION  303216 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION   BG384217
VERSION     BG384217.1  GI:13308689
KEYWORDS    EST.
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 535)
  AUTHORS   Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
            Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
            and Keele,J.W.
  TITLE     Design and use of two pooled tissue normalized cDNA libraries for
            EST discovery in swine
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 90  row: G  column: 13
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
                     /clone_lib="MARC 1PIG"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                     Library made from pooled tissue from day 11, 13, 15, 20,
                     and 30 embryos."
BASE COUNT      121 a    159 c    136 g    119 t
ORIGIN


  Query Match           13.6%;  Score 30;  DB 10;  Length 535;
  Best Local Similarity  100.0%;  Pred. No. 0.00012;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    192 aagagacgcaaacacaaaagtggaaaacag 221
          ||||||||||||||||||||||||||||||
Db    151 AAGAGACGCAAACACAAAAGTGGAAAACAG 180


RESULT   6
AG037352/c
LOCUS       AG037352                679 bp    DNA     linear   GSS 01-NOV-2001
DEFINITION  Pan troglodytes DNA, clone: PTB-013J06.R, genomic survey sequence.
ACCESSION   AG037352
VERSION     AG037352.1  GI:16564225
KEYWORDS    GSS; GSS (genome survey sequence).
SOURCE      Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male
            BAC Library clone:PTB-013J06.R.
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (sites)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library PTB
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 679)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
```

```
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library PTB This BAC end
            was generated during the R&D process and may have higher chance of
            clone tracking errors.
             PRIMERS
                 Sequencing: M13Rev
             LIBRARY
                 Vector    : pKS145
                 R.Site 1  : SacI
                 R.Site 2  : SacI.
FEATURES                 Location/Qualifiers
     source              1. .679
                         /organism="Pan troglodytes"
                         /db_xref="taxon:9598"
                         /clone="PTB-013J06.R"
                         /sex="male"
                         /cell_type="lymphoblast"
                         /clone_lib="PTB Chimpanzee Male BAC Library"
BASE COUNT       357 a     89 c    201 g     22 t     10 others
ORIGIN

  Query Match                 9.5%;  Score 21;  DB 12;  Length 679;
  Best Local Similarity   100.0%;  Pred. No. 4.7;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     146 ttctgttttctccccttctcc 166
           |||||||||||||||||||||
Db     431 TTCTGTTTTCTCCCCTTCTCC 411


RESULT   7
BH502182
LOCUS       BH502182                710 bp    DNA     linear   GSS 13-DEC-2001
DEFINITION  BOGHN04TR BOGH Brassica oleracea genomic clone BOGHN04, DNA
            sequence.
ACCESSION   BH502182
VERSION     BH502182.1  GI:17710279
KEYWORDS    GSS.
SOURCE      Brassica oleracea.
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 710)
  AUTHORS   Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
  TITLE     Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL   Unpublished (2001)
COMMENT     Other_GSSs: BOGHN04TF
            Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TR
            Class: sheared ends.
FEATURES                 Location/Qualifiers
     source              1. .710
                         /organism="Brassica oleracea"
                         /strain="TO1000DH3"
                         /db_xref="taxon:3712"
                         /clone="BOGHN04"
                         /clone_lib="BOGH"
                         /note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
                         genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT       177 a    196 c    128 g    209 t
ORIGIN

  Query Match                 9.5%;  Score 21;  DB 12;  Length 710;
  Best Local Similarity   100.0%;  Pred. No. 4.7;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     134 ttttttccccggttctgtttt 154
           |||||||||||||||||||||
Db     532 TTTTTTCCCCGGTTCTGTTTT 552


RESULT   8
BB862540
LOCUS       BB862540                265 bp    mRNA    linear   EST 26-NOV-2001
DEFINITION  BB862540 RIKEN full-length enriched, brain CRL-1443 BC3H1 cDNA Mus
            musculus cDNA clone G430028E10 5', mRNA sequence.
ACCESSION   BB862540
VERSION     BB862540.1  GI:17103994
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 265)
  AUTHORS   Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
            Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii
            ,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
            Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.
            , Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
            Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
            ,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.
            , Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
            2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES                 Location/Qualifiers
     source              1. .265
                         /organism="Mus musculus"
                         /strain="C3H"
                         /db_xref="taxon:10090"
                         /clone="G430028E10"
                         /clone_lib="RIKEN full-length enriched, brain CRL-1443
                         BC3H1 cDNA"
                         /tissue_type="brain"
                         /cell_line="CRL-1443 BC3H1"
BASE COUNT        45 a     73 c    114 g     33 t
ORIGIN
```

```
  Query Match             9.0%;  Score 20;  DB 9;  Length 265;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      68 gcggcagggcagggcgggga 87
           ||||||||||||||||||||
Db     148 GCGGCAGGGCAGGGCGGGGA 167


RESULT   9
BG629097
LOCUS       BG629097                389 bp    mRNA    linear   EST 19-APR-2001
DEFINITION  cC-esflcLEL25O23d1 Tomato flower library from a mixture of
            developmental stages Lycopersicon esculentum cDNA clone
            cC-esflcLEL25O23d1, mRNA sequence.
ACCESSION   BG629097
VERSION     BG629097.1  GI:13680570
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 389)
  AUTHORS   van der Hoeven,R.S. and Tanksley,S.D.
  TITLE     ESTs from a tomato flower library
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Rutger S. van der Hoeven
            Cornell University
            252 Emerson Hall, Ithaca, NY 14850, USA
            Tel: 607 255 7886
            Fax: 607 255 6683
            Email: rv19@cornell.edu
            3 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .389
                     /organism="Lycopersicon esculentum"
                     /cultivar="E6203"
                     /db_xref="taxon:4081"
                     /clone="cC-esflcLEL25O23d1"
                     /clone_lib="Tomato flower library from a mixture of
                     developmental stages"
                     /tissue_type="developing flower buds and open flowers"
                     /dev_stage="4-8 week old plants"
                     /lab_host="XLOLR"
                     /note="Vector: pBK_CMV; Site_1: EcoRI; Site_2: XhoI;
                     Flowers and flower buds were collected from greenhouse
                     grown plants  and used for library construction (cLEL). "
BASE COUNT       68 a    100 c     71 g    150 t
ORIGIN


  Query Match             9.0%;  Score 20;  DB 10;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     124 ttgctccttgtttttttcccc 143
           ||||||||||||||||||||
Db     246 TTGCTCCTTGTTTTTTCCCC 265


RESULT  10
BG629091
LOCUS       BG629091                426 bp    mRNA    linear   EST 19-APR-2001
DEFINITION  cC-esflcLEL25O11a1 Tomato flower library from a mixture of
            developmental stages Lycopersicon esculentum cDNA clone
            cC-esflcLEL25O11a1, mRNA sequence.
ACCESSION   BG629091
VERSION     BG629091.1  GI:13680564
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 426)
  AUTHORS   van der Hoeven,R.S. and Tanksley,S.D.
  TITLE     ESTs from a tomato flower library
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Rutger S. van der Hoeven
            Cornell University
            252 Emerson Hall, Ithaca, NY 14850, USA
            Tel: 607 255 7886
            Fax: 607 255 6683
            Email: rv19@cornell.edu
            3 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .426
                     /organism="Lycopersicon esculentum"
                     /cultivar="E6203"
                     /db_xref="taxon:4081"
                     /clone="cC-esflcLEL25O11a1"
                     /clone_lib="Tomato flower library from a mixture of
                     developmental stages"
                     /tissue_type="developing flower buds and open flowers"
                     /dev_stage="4-8 week old plants"
                     /lab_host="XLOLR"
                     /note="Vector: pBK_CMV; Site_1: EcoRI; Site_2: XhoI;
                     Flowers and flower buds were collected from greenhouse
                     grown plants  and used for library construction (cLEL). "
BASE COUNT       77 a    108 c     79 g    162 t
ORIGIN


  Query Match             9.0%;  Score 20;  DB 10;  Length 426;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     124 ttgctccttgtttttttcccc 143
           ||||||||||||||||||||
Db     246 TTGCTCCTTGTTTTTTCCCC 265


RESULT  11
BG123629/c
LOCUS       BG123629                777 bp    mRNA    linear   EST 31-JAN-2001
DEFINITION  EST469275 tomato shoot/meristem Lycopersicon esculentum cDNA clone
            cTOF2H15 5' sequence, mRNA sequence.
ACCESSION   BG123629
VERSION     BG123629.1  GI:12623817
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 777)
  AUTHORS   van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Utterback,T.,
            Hansen,C., Ronning,C. and Tanksley,S.
  TITLE     Generation of ESTs from tomato shoot/meristem tissue
  JOURNAL   Unpublished (2001)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html.
FEATURES             Location/Qualifiers
     source          1. .777
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
```

```
                     /clone="cTOF2H15"
                     /clone_lib="tomato shoot/meristem"
                     /tissue_type="shoot/meristem"
                     /dev_stage="developing shoots from 4-6wks old plants"
                     /lab_host="SOLR"
                     /note="Vector: pBluescript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; Small expanding leaves from the growing tip were
                     taken from greenhouse plants (4-6wks old TA496). Tissue
                     was immediately frozen in liquid nitrogen."
BASE COUNT       262 a    143 c    210 g    162 t
ORIGIN


  Query Match             9.0%;  Score 20;  DB 10;  Length 777;
  Best Local Similarity   100.0%;  Pred. No. 15;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    124 ttgctccttgttttttcccc 143
          ||||||||||||||||||||
Db    495 TTGCTCCTTGTTTTTTCCCC 476


RESULT  12
AU076672/c
LOCUS       AU076672                 87 bp    mRNA    linear   EST 04-MAY-2000
DEFINITION  AU076672 Sugano cDNA library Homo sapiens cDNA clone HEP02824
            similar to 5'-end region of Human interferon-gamma receptor mRNA,
            mRNA sequence.
ACCESSION   AU076672
VERSION     AU076672.1  GI:7439153
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 87)
  AUTHORS   Suzuki,Y., Ishihara,D., Sasaki,M., Nakagawa,H., Hata,H., Tsunoda,T.
            , Watanabe,M., Komatsu,T., Ota,T., Isogai,T., Suyama,A. and Sugano
            ,S.
  TITLE     Statistical analysis of the 5' untranslated region of human mRNA
            using 'Oligo-Capped' cDNA libraries
  JOURNAL   Genomics 64 (3), 286-297 (2000)
  MEDLINE   20221373
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997)
             This clone was obtained from a 'full length-enriched' cDNA library
            constructed by 'Oligo-Capping' method. The coding region starts
            from the 50 bp upstream to the 3'-end.
FEATURES             Location/Qualifiers
     source          1. .87
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HEP02824"
                     /clone_lib="Sugano cDNA library"
BASE COUNT       11 a     27 c     28 g     21 t
ORIGIN


  Query Match             8.6%;  Score 19;  DB 9;  Length 87;
  Best Local Similarity   100.0%;  Pred. No. 58;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    178 caaggggtaggagaaagag 196
          ||||||||||||||||||||
Db     65 CAAGGGGTAGGAGAAAGAG 47


RESULT  13
AU099072/c
LOCUS       AU099072                 300 bp    mRNA    linear   EST 05-APR-2001
DEFINITION  AU099072 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HEP16514 similar to Human interferon-gamma receptor mRNA, mRNA
            sequence.
ACCESSION   AU099072
VERSION     AU099072.1  GI:13550201
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 300)
  AUTHORS   Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
            ,K., Suyama,A. and Sugano,S.
  TITLE     In silico mapping of the 5'-ends of human mRNAs using full-length
            enriched and 5'-end enriched cDNA libraries constructed by
            Oligo-capping method
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .300
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HEP16514"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT       62 a     76 c     87 g     75 t
ORIGIN


  Query Match             8.6%;  Score 19;  DB 9;  Length 300;
  Best Local Similarity   100.0%;  Pred. No. 52;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    178 caaggggtaggagaaagag 196
          ||||||||||||||||||||
Db    116 CAAGGGGTAGGAGAAAGAG 98


RESULT  14
BI772330/c
LOCUS       BI772330                 389 bp    mRNA    linear   EST 25-SEP-2001
DEFINITION  603056038F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5205633 5',
            mRNA sequence.
ACCESSION   BI772330
VERSION     BI772330.1  GI:15763908
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 389)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
```

```
          Plate: LLAM11515  row: k  column: 10
          High quality sequence stop: 389.
FEATURES             Location/Qualifiers
     source          1. .389
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5205633"
                     /clone_lib="NIH_MGC_122"
                     /lab_host="DH10B"
                     /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
                     Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
                     anonymous pool of 24 week female lung, 16 week female
                     spleen, and 20-22 week male spleens.  Library is oligo-dT
                     primed and directionally cloned (EcoRV site is destroyed
                     upon  cloning). Average insert size 1.4 kb, insert size
                     range 1-3 kb. Library is normalized and enriched for
                     full-length clones and was  constructed by C. Gruber
                     (Invitrogen).  Research Genetics tracking  code 026. Note:
                     this is a NIH_MGC Library."
BASE COUNT       96 a     91 c     93 g    109 t
ORIGIN


  Query Match             8.6%;  Score 19;  DB 10;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 51;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    178 caaggggtaggagaaagag 196
          ||||||||||||||||||||
Db     60 CAAGGGGTAGGAGAAAGAG 42


RESULT  15
AW642780
LOCUS       AW642780                 484 bp    mRNA    linear   EST 26-APR-2001
DEFINITION  cm22e02.w1 Blackshear/Soares normalized Xenopus egg library Xenopus
            laevis cDNA clone PBX0121E02 5', mRNA sequence.
ACCESSION   AW642780
VERSION     AW642780.1  GI:7400093
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 484)
  AUTHORS   Blackshear,P.J., Lai,W.S., Thorn,J.M., Kennington,E.A., Staffa,N.G.
            Jr., Moore,D.T., Bouffard,G.G., Beckstrom-Sternberg,S.M., Touchman
            ,J.W., Bonaldo,M.F. and Soares,M.B.
  TITLE     The NIEHS Xenopus maternal EST project: interim analysis of the
            first 13,879 ESTs from unfertilized eggs
  JOURNAL   Gene 267 (1), 71-87 (2001)
  MEDLINE   21211403
COMMENT     Contact: Perry J. Blackshear
            Office of Clinical Research and Laboratory of Signal Transduction
            National Institute of Environmental Health Sciences
            A2-05 NIEHS, 101 Alexander Drive, Research Triangle Park, NC 27709,
            USA
            Tel: 919 541-4899
            Fax: 919 541-4571
            Email: black009@niehs.nih.gov
            Clone is available through Research Genetics, Inc., 2130 Memorial
            Parkway, Huntsville, AL 35901
             phone 800-533-4363 ext.cdna, fax 256-536-9016 att:cdna, email
            cdna@resgen.com
             DNA Sequencing and analyses performed by National Institutes of
            Health Intramural Sequencing Center (NISC).
            PCR PRimers
            FORWARD: TGTAAAACGACGGCCAGT
            BACKWARD: CAGGAAACAGCTATGACC
            Plate: 0121  row: E  column: 02
            Seq primer: T7 primer.
FEATURES             Location/Qualifiers
     source          1. .484
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="PBX0121E02"
                     /clone_lib="Blackshear/Soares normalized Xenopus egg
                     library"
                     /sex="female"
                     /tissue_type="unfertilized egg"
                     /cell_type="unfertilized egg"
                     /dev_stage="unfertilized egg"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3-Pac; Site_1: EcoRI; Site_2: NotI;
                     PolyA-selected mRNA was prepared from unfertilized Xenopus
                     laevis eggs.  The library was constructed in the vector
                     pT7T3-Pac as described in Bonaldo, M.F., Lennon, G. and
                     Soares, M.B. 'Normalization and subtraction:  two
                     approaches to facilitate gene discovery',  Genome Research
                     6:791-806, 1996.  The first strand synthesis used a
                     NotI-dT18 primer; double stranded cDNAs were ligated to
                     EcoRI adapters, digested with NotI, and directionally
                     cloned into the NotI and EcoRI-digested pT7T3-Pac vector.
                     The library contained approximately 7.2 X 10^5
                     recombinants, with average insert sizes of 1-1.5 kb."
BASE COUNT      120 a     83 c    110 g    171 t
ORIGIN


  Query Match             8.6%;  Score 19;  DB 9;  Length 484;
  Best Local Similarity   100.0%;  Pred. No. 50;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    113 ctcagggcgctttgctcct 131
          ||||||||||||||||||||
Db    345 CTCAGGGCGCTTTGCTCCT 363


RESULT  16
T78163
LOCUS       T78163                   486 bp    mRNA    linear   EST 15-MAR-1995
DEFINITION  yd79b06.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:114419 5' similar to gb:J03143 INTERFERON-GAMMA RECEPTOR
            ALPHA CHAIN PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION   T78163
VERSION     T78163.1  GI:696672
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 486)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Other_ESTs: yd79b06.s1
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1053
            High quality sequence stops: 475 Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1053   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 475.
FEATURES             Location/Qualifiers
     source          1. .486
                     /organism="Homo sapiens"
```

```
                     /db_xref="GDB:470036"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:114419"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT       80 a    134 c    149 g    101 t     22 others
ORIGIN


  Query Match             8.6%;  Score 19;  DB 10;  Length 486;
  Best Local Similarity   100.0%;  Pred. No. 50;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    178 caaggggtaggagaaagag 196
          ||||||||||||||||||||
Db    129 CAAGGGGTAGGAGAAAGAG 147


RESULT  17
BG702824/c
LOCUS       BG702824                530 bp    mRNA    linear   EST 07-MAY-2001
DEFINITION  602684727F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4817263 5',
            mRNA sequence.
ACCESSION   BG702824
VERSION     BG702824.1  GI:13974552
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 530)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10717  row: e  column: 08
            High quality sequence stop: 474.
FEATURES             Location/Qualifiers
     source          1. .530
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4817263"
                     /clone_lib="NIH_MGC_95"
                     /tissue_type="hippocampus"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.5 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for  full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health).  Note: this is a NIH_MGC Library."
BASE COUNT      154 a    110 c    132 g    134 t
ORIGIN


  Query Match             8.6%;  Score 19;  DB 10;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 50;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    178 caaggggtaggagaaagag 196
          ||||||||||||||||||||
Db     64 CAAGGGGTAGGAGAAAGAG 46


RESULT  18
AL597062/c
LOCUS       AL597062                554 bp    mRNA    linear   EST 14-AUG-2001
DEFINITION  DKFZp313C0311_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone
            DKFZp313C0311 5', mRNA sequence.
ACCESSION   AL597062
VERSION     AL597062.1  GI:15154758
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 554)
  AUTHORS   Koehrer,K., Beyer,A., Mewes,W., Weil,B. and Wiemann,S.
  TITLE     EST (Koehrer,K., Beyer,A., Mewes,H.W., Weil,B. and Wiemann,S.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Koehrer K
            MIPS
            Am Klopferspitz 18a D-82152 Martinsried, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by BMFZ (Biomedical Research Center at the Charite,
            Berlin/Germany) within the cDNA sequencing consortium of the German
            Genome Project.
            No s1 sequence available.
            This clone (DKFZp313C0311) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .554
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp313C0311"
                     /clone_lib="313 (synonym: hlcc2)"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
BASE COUNT      150 a    119 c    105 g    180 t
ORIGIN


  Query Match             8.6%;  Score 19;  DB 9;  Length 554;
  Best Local Similarity   100.0%;  Pred. No. 50;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    178 caaggggtaggagaaagag 196
          ||||||||||||||||||||
Db     60 CAAGGGGTAGGAGAAAGAG 42


RESULT  19
BI548754/c
LOCUS       BI548754                597 bp    mRNA    linear   EST 05-SEP-2001
DEFINITION  603189147F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5260619 5',
            mRNA sequence.
ACCESSION   BI548754
VERSION     BI548754.1  GI:15436066
```

```
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 597)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11656  row: n  column: 12
            High quality sequence stop: 593.
FEATURES             Location/Qualifiers
     source          1. .597
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5260619"
                     /clone_lib="NIH_MGC_95"
                     /tissue_type="hippocampus"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.5 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for  full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health).  Note: this is a NIH_MGC Library."
BASE COUNT      163 a    124 c    150 g    160 t
ORIGIN


  Query Match             8.6%;  Score 19;  DB 10;  Length 597;
  Best Local Similarity   100.0%;  Pred. No. 49;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    178 caaggggtaggagaaagag 196
          ||||||||||||||||||
Db     91 CAAGGGGTAGGAGAAAGAG 73


RESULT  20
BI523553/c
LOCUS       BI523553                609 bp    mRNA    linear   EST 29-AUG-2001
DEFINITION  603175746F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5240195 5',
            mRNA sequence.
ACCESSION   BI523553
VERSION     BI523553.1  GI:15348345
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 609)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11605  row: k  column: 12
            High quality sequence start: 31
            High quality sequence stop: 608.
FEATURES             Location/Qualifiers
     source          1. .609
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5240195"
                     /clone_lib="NIH_MGC_121"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: EcoRV (destroyed); RNA source anonymous pool of 3
                     fetal brains, female age 20 weeks, female age 24 weeks,
                     and male age 26 weeks.  Library is oligo-dT primed and
                     directionally cloned (EcoRV site is destroyed  upon
                     cloning).  Average insert size 1.7 kb, insert size range
                     0.7-3.5 kb.  Library is normalized and enriched for
                     full-length clones and  was constructed by C. Gruber
                     (Invitrogen).  Research Genetics  tracking code 017. Note:
                     this is a NIH_MGC Library."
BASE COUNT      184 a    120 c    143 g    162 t
ORIGIN


  Query Match             8.6%;  Score 19;  DB 10;  Length 609;
  Best Local Similarity   100.0%;  Pred. No. 49;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    178 caaggggtaggagaaagag 196
          ||||||||||||||||||
Db     59 CAAGGGGTAGGAGAAAGAG 41


RESULT  21
BG714492/c
LOCUS       BG714492                616 bp    mRNA    linear   EST 08-MAY-2001
DEFINITION  602670936F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4793528 5',
            mRNA sequence.
ACCESSION   BG714492
VERSION     BG714492.1  GI:13993423
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 616)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10673  row: h  column: 09
            High quality sequence stop: 616.
FEATURES             Location/Qualifiers
     source          1. .616
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4793528"
                     /clone_lib="NIH_MGC_96"
                     /tissue_type="hypothalamus"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pBluescriptR (modified
```

pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 182 a 123 c 150 g 161 t
ORIGIN

Query Match 8.6%; Score 19; DB 10; Length 616;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 178 caaggggtaggagaaagag 196
||||||||||||||||||||
Db 68 CAAGGGGTAGGAGAAAGAG 50

RESULT 22
BG776309/c
LOCUS BG776309 656 bp mRNA linear EST 15-MAY-2001
DEFINITION 602663377F1 NIH_MGC_59 Homo sapiens cDNA clone IMAGE:4808562 5', mRNA sequence.
ACCESSION BG776309
VERSION BG776309.1 GI:14046626
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 656)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLCM1661 row: j column: 19
High quality sequence stop: 656.
FEATURES Location/Qualifiers
source 1. .656
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4808562"
/clone_lib="NIH_MGC_59"
/tissue_type="mucoepidermoid carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."
BASE COUNT 189 a 134 c 160 g 173 t
ORIGIN

Query Match 8.6%; Score 19; DB 10; Length 656;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 178 caaggggtaggagaaagag 196
||||||||||||||||||||
Db 60 CAAGGGGTAGGAGAAAGAG 42

RESULT 23
BM311005/c
LOCUS BM311005 664 bp mRNA linear EST 03-JAN-2002
DEFINITION ig59e10.y1 HR85 islet Homo sapiens cDNA 5' similar to SW:INGR_HUMAN P15260 INTERFERON-GAMMA RECEPTOR ALPHA CHAIN PRECURSOR ;, mRNA sequence.
ACCESSION BM311005
VERSION BM311005.1 GI:18045350
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 664)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T. , Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 485.
FEATURES Location/Qualifiers
source 1. .664
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HR85 islet"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1: NotI; Site_2: XhoI; cDNA made by oligo-dT priming. Size-selected on agarose gel. Average insert size -1kb. 5' XhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permutt Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel: 314-362-1916, Fax: 314-747-2692."
BASE COUNT 194 a 134 c 157 g 179 t
ORIGIN

Query Match 8.6%; Score 19; DB 10; Length 664;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 178 caaggggtaggagaaagag 196
||||||||||||||||||||
Db 53 CAAGGGGTAGGAGAAAGAG 35

```
RESULT  24
BI561894/c
LOCUS       BI561894                674 bp    mRNA    linear   EST 05-SEP-2001
DEFINITION  603255765F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5297945 5',
            mRNA sequence.
ACCESSION   BI561894
VERSION     BI561894.1  GI:15449208
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 674)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11754  row: a  column: 18
            High quality sequence stop: 674.
FEATURES             Location/Qualifiers
     source          1. .674
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5297945"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      192 a    139 c    166 g    177 t
ORIGIN


  Query Match             8.6%;  Score 19;  DB 10;  Length 674;
  Best Local Similarity   100.0%;  Pred. No. 49;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     178 caaggggtaggagaaagag 196
           ||||||||||||||||||||
Db      74 CAAGGGGTAGGAGAAAGAG 56


RESULT  25
BI545327/c
LOCUS       BI545327                686 bp    mRNA    linear   EST 05-SEP-2001
DEFINITION  603187458F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5258823 5',
            mRNA sequence.
ACCESSION   BI545327
VERSION     BI545327.1  GI:15432639
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 686)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11652  row: c  column: 16
            High quality sequence stop: 568.
FEATURES             Location/Qualifiers
     source          1. .686
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5258823"
                     /clone_lib="NIH_MGC_95"
                     /tissue_type="hippocampus"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.5 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for  full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health).  Note: this is a NIH_MGC Library."
BASE COUNT      188 a    152 c    170 g    176 t
ORIGIN


  Query Match             8.6%;  Score 19;  DB 10;  Length 686;
  Best Local Similarity   100.0%;  Pred. No. 49;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     178 caaggggtaggagaaagag 196
           ||||||||||||||||||||
Db      91 CAAGGGGTAGGAGAAAGAG 73


RESULT  26
BI549570
LOCUS       BI549570                688 bp    mRNA    linear   EST 05-SEP-2001
DEFINITION  603192272F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5263721 5',
            mRNA sequence.
ACCESSION   BI549570
VERSION     BI549570.1  GI:15436882
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 688)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11664  row: o  column: 18
            High quality sequence stop: 688.
FEATURES             Location/Qualifiers
     source          1. .688
                     /organism="Homo sapiens"
```

/db_xref="taxon:9606"
/clone="IMAGE:5263721"
/clone_lib="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.5 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 149 a 221 c 218 g 100 t
ORIGIN

Query Match 8.6%; Score 19; DB 10; Length 688;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 178 caaggggtaggagaaagag 196
||||||||||||||||||||
Db 456 CAAGGGGTAGGAGAAAGAG 474

RESULT 27
BJ044199/c
LOCUS BJ044199 694 bp mRNA linear EST 06-DEC-2001
DEFINITION BJ044199 NIBB Mochii normalized Xenopus neurula library Xenopus laevis cDNA clone XL012i22 3', mRNA sequence.
ACCESSION BJ044199
VERSION BJ044199.1 GI:17397688
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 694)
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara ,Y.
TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES Location/Qualifiers
source 1. .694
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL012i22"
/clone_lib="NIBB Mochii normalized Xenopus neurula library"
/tissue_type="whole embryo"
/dev_stage="stage 15"
BASE COUNT 230 a 129 c 119 g 216 t
ORIGIN

Query Match 8.6%; Score 19; DB 10; Length 694;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 113 ctcagggcgctttgctcct 131
||||||||||||||||||||
Db 632 CTCAGGGCGCTTTGCTCCT 614

RESULT 28
BJ052576/c
LOCUS BJ052576 703 bp mRNA linear EST 11-DEC-2001
DEFINITION BJ052576 NIBB Mochii normalized Xenopus neurula library Xenopus laevis cDNA clone XL041o19 3', mRNA sequence.
ACCESSION BJ052576
VERSION BJ052576.1 GI:17498622
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 703)
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara ,Y.
TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES Location/Qualifiers
source 1. .703
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL041o19"
/clone_lib="NIBB Mochii normalized Xenopus neurula library"
/tissue_type="whole embryo"
/dev_stage="stage 15"
BASE COUNT 235 a 129 c 122 g 217 t
ORIGIN

Query Match 8.6%; Score 19; DB 10; Length 703;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 113 ctcagggcgctttgctcct 131
||||||||||||||||||||
Db 644 CTCAGGGCGCTTTGCTCCT 626

RESULT 29
AL554654/c
LOCUS AL554654 704 bp mRNA linear EST 16-FEB-2001
DEFINITION AL554654 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI085YJ06 5 prime, mRNA sequence.
ACCESSION AL554654
VERSION AL554654.1 GI:12895644
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 704)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES Location/Qualifiers
source 1. .704
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI085YJ06"

/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
BASE COUNT 204 a 132 c 169 g 191 t 8 others
ORIGIN

Query Match 8.6%; Score 19; DB 9; Length 704;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 178 caaggggtaggagaaagag 196
||||||||||||||||||||
Db 59 CAAGGGGTAGGAGAAAGAG 41

RESULT 30
BG431399/c
LOCUS BG431399 706 bp mRNA linear EST 14-MAR-2001
DEFINITION 602500027F1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4613742 5', mRNA sequence.
ACCESSION BG431399
VERSION BG431399.1 GI:13337905
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 706)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLCM1364 row: e column: 07
High quality sequence stop: 701.
FEATURES Location/Qualifiers
source 1. .706
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4613742"
/clone_lib="NIH_MGC_75"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."
BASE COUNT 212 a 131 c 166 g 197 t
ORIGIN

Query Match 8.6%; Score 19; DB 10; Length 706;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 178 caaggggtaggagaaagag 196
||||||||||||||||||||
Db 26 CAAGGGGTAGGAGAAAGAG 8

RESULT 31
AL600963/c
LOCUS AL600963 725 bp mRNA linear EST 14-AUG-2001
DEFINITION DKFZp313B0639_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone DKFZp313B0639 5', mRNA sequence.
ACCESSION AL600963
VERSION AL600963.1 GI:15164469
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 725)
AUTHORS Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE EST (Duesterhoeft, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Duesterhoeft A
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de; sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing consortium of the German Genome Project.
No s1 sequence available.
This clone (DKFZp313B0639) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES Location/Qualifiers
source 1. .725
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp313B0639"
/clone_lib="313 (synonym: hlcc2)"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB; cDNA-collection"
BASE COUNT 206 a 141 c 174 g 193 t 11 others
ORIGIN

Query Match 8.6%; Score 19; DB 9; Length 725;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 178 caaggggtaggagaaagag 196
||||||||||||||||||||
Db 62 CAAGGGGTAGGAGAAAGAG 44

RESULT 32
BE973918/c
LOCUS BE973918 730 bp mRNA linear EST 04-OCT-2000
DEFINITION 601680234F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:3950528 5', mRNA sequence.
ACCESSION BE973918
VERSION BE973918.1 GI:10587254
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```
REFERENCE   1  (bases 1 to 730)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONETECH Laboratories, Inc.
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM817  row: c  column: 09
            High quality sequence stop: 577.
FEATURES             Location/Qualifiers
     source          1. .730
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3950528"
                     /clone_lib="NIH_MGC_83"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: prostate; Vector: pDNR-LIB (Clontech);
                     Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc
                     ); 5' and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).   Average
                     insert size 1.4 kb (range 0.5-4.0 kb).  14/15 colonies
                     contained inserts  by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA)."
BASE COUNT      203 a    155 c    184 g    188 t
ORIGIN


  Query Match             8.6%;  Score 19;  DB 10;  Length 730;
  Best Local Similarity   100.0%;  Pred. No. 49;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    178 caaggggtaggagaaagag 196
          ||||||||||||||||||
Db     91 CAAGGGGTAGGAGAAAGAG 73


RESULT  33
BI824845/c
LOCUS       BI824845                731 bp    mRNA    linear   EST 04-OCT-2001
DEFINITION  603033758F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5174751 5',
            mRNA sequence.
ACCESSION   BI824845
VERSION     BI824845.1  GI:15936395
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 731)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11435  row: d  column: 16
            High quality sequence stop: 729.
FEATURES             Location/Qualifiers
     source          1. .731
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5174751"
                     /clone_lib="NIH_MGC_115"
                     /lab_host="DH10B"
                     /note="Organ: pooled brain, lung, testis; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 6 male brains, age range 23-27; 1
                     male lung, age 27; and 1 male testis, age 69.  Library is
                     oligo-dT primed and directionally cloned (EcoRV site is
                     destroyed upon  cloning).  Average insert size 1.8 kb,
                     insert size range 1-3 kb. Library is normalized and
                     enriched for full-length clones and was  constructed by C.
                     Gruber (Invitrogen).  Research Genetics tracking  code
                     021. Note: this is a NIH_MGC Library."
BASE COUNT      216 a    143 c    173 g    199 t
ORIGIN


  Query Match             8.6%;  Score 19;  DB 10;  Length 731;
  Best Local Similarity   100.0%;  Pred. No. 49;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    178 caaggggtaggagaaagag 196
          ||||||||||||||||||
Db     50 CAAGGGGTAGGAGAAAGAG 32


RESULT  34
BG564438/c
LOCUS       BG564438                748 bp    mRNA    linear   EST 10-APR-2001
DEFINITION  602584384F1 NIH_MGC_76 Homo sapiens cDNA clone IMAGE:4712109 5',
            mRNA sequence.
ACCESSION   BG564438
VERSION     BG564438.1  GI:13572090
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 748)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONETECH Laboratories, Inc.
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1553  row: g  column: 22
            High quality sequence stop: 709.
FEATURES             Location/Qualifiers
     source          1. .748
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4712109"
                     /clone_lib="NIH_MGC_76"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: liver; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).  Average insert size 1.85
                     kb (range 1.0-4.0 kb). 15/15 colonies contained  inserts
                     by PCR. This library was enriched for full-length clones
                     and was constructed by Clontech  Laboratories (Palo Alto,
                     CA). Note: this is a NIH_MGC Library."
```

```
BASE COUNT      215 a    146 c    184 g    203 t
ORIGIN


  Query Match             8.6%;  Score 19;  DB 10;  Length 748;
  Best Local Similarity   100.0%;  Pred. No. 49;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    178 caaggggtaggagaaagag 196
          ||||||||||||||||||||
Db     60 CAAGGGGTAGGAGAAAGAG 42


RESULT  35
AL601135/c
LOCUS       AL601135                751 bp    mRNA    linear   EST 14-AUG-2001
DEFINITION  DKFZp313O2239_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone
            DKFZp313O2239 5', mRNA sequence.
ACCESSION   AL601135
VERSION     AL601135.1  GI:15164641
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 751)
  AUTHORS   Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann
            ,S.
  TITLE     EST (Duesterhoeft, et al.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Duesterhoeft A
            MIPS
            Am Klopferspitz 18a D-82152 Martinsried, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            No s1 sequence available.
            This clone (DKFZp313O2239) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .751
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp313O2239"
                     /clone_lib="313 (synonym: hlcc2)"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
BASE COUNT      207 a    147 c    188 g    203 t      6 others
ORIGIN


  Query Match             8.6%;  Score 19;  DB 9;  Length 751;
  Best Local Similarity   100.0%;  Pred. No. 49;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    178 caaggggtaggagaaagag 196
          ||||||||||||||||||||
Db     87 CAAGGGGTAGGAGAAAGAG 69


RESULT  36
BI088379/c
LOCUS       BI088379                756 bp    mRNA    linear   EST 20-JUN-2001
DEFINITION  602851164F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:4992998 5',
            mRNA sequence.
ACCESSION   BI088379
VERSION     BI088379.1  GI:14506709
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 756)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Incyte Genomics, Inc.
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11012  row: o  column: 15
            High quality sequence stop: 660.
FEATURES             Location/Qualifiers
     source          1. .756
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4992998"
                     /clone_lib="NIH_MGC_10"
                     /cell_line="MGC36"
                     /lab_host="DH10B"
                     /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.5 kb. Library prepared by Life
                     Technologies."
BASE COUNT      216 a    154 c    193 g    193 t
ORIGIN


  Query Match             8.6%;  Score 19;  DB 10;  Length 756;
  Best Local Similarity   100.0%;  Pred. No. 49;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    178 caaggggtaggagaaagag 196
          ||||||||||||||||||||
Db    107 CAAGGGGTAGGAGAAAGAG 89


RESULT  37
BG706900/c
LOCUS       BG706900                757 bp    mRNA    linear   EST 07-MAY-2001
DEFINITION  602672070F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4794785 5',
            mRNA sequence.
ACCESSION   BG706900
VERSION     BG706900.1  GI:13982706
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 757)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10676  row: 1  column: 18
            High quality sequence stop: 757.
FEATURES             Location/Qualifiers
```

```
     source          1. .757
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4794785"
                     /clone_lib="NIH_MGC_96"
                     /tissue_type="hypothalamus"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.3 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length  clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      206 a    157 c    197 g    197 t
ORIGIN


  Query Match             8.6%;  Score 19;  DB 10;  Length 757;
  Best Local Similarity   100.0%;  Pred. No. 49;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     178 caaggggtaggagaaagag 196
           ||||||||||||||||||||
Db     111 CAAGGGGTAGGAGAAAGAG 93


RESULT  38
AU143746/c
LOCUS       AU143746                781 bp    mRNA    linear   EST 25-OCT-2000
DEFINITION  AU143746 Y79AA1 Homo sapiens cDNA clone Y79AA1002429 5', mRNA
            sequence.
ACCESSION   AU143746
VERSION     AU143746.1  GI:11005267
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 781)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .781
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Y79AA1002429"
                     /clone_lib="Y79AA1"
                     /cell_type="retinoblastoma"
                     /cell_line="Y79"
                     /note="Vector: pME18SFL3"
BASE COUNT      225 a    156 c    190 g    207 t      3 others
ORIGIN


  Query Match             8.6%;  Score 19;  DB 9;  Length 781;
  Best Local Similarity   100.0%;  Pred. No. 48;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     178 caaggggtaggagaaagag 196
           ||||||||||||||||||||
Db      86 CAAGGGGTAGGAGAAAGAG 68


RESULT  39
BI458234/c
LOCUS       BI458234                784 bp    mRNA    linear   EST 21-AUG-2001
DEFINITION  603199174F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5278435 5',
            mRNA sequence.
ACCESSION   BI458234
VERSION     BI458234.1  GI:15248890
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 784)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11703  row: d  column: 20
            High quality sequence stop: 780.
FEATURES             Location/Qualifiers
     source          1. .784
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5278435"
                     /clone_lib="NIH_MGC_96"
                     /tissue_type="hypothalamus"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.3 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length  clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      214 a    159 c    212 g    198 t      1 others
ORIGIN


  Query Match             8.6%;  Score 19;  DB 10;  Length 784;
  Best Local Similarity   100.0%;  Pred. No. 48;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     178 caaggggtaggagaaagag 196
           ||||||||||||||||||||
Db     136 CAAGGGGTAGGAGAAAGAG 118


RESULT  40
BG705539/c
LOCUS       BG705539                795 bp    mRNA    linear   EST 07-MAY-2001
DEFINITION  602668716F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4791771 5',
            mRNA sequence.
ACCESSION   BG705539
VERSION     BG705539.1  GI:13979982
KEYWORDS    EST.
```

```
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 795)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10668  row: o  column: 04
            High quality sequence stop: 675.
FEATURES             Location/Qualifiers
     source          1. .795
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4791771"
                     /clone_lib="NIH_MGC_96"
                     /tissue_type="hypothalamus"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.3 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length  clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      223 a    172 c    188 g    212 t
ORIGIN


  Query Match             8.6%;  Score 19;  DB 10;  Length 795;
  Best Local Similarity   100.0%;  Pred. No. 48;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    178 caaggggtaggagaaagag 196
          ||||||||||||||||||
Db     68 CAAGGGGTAGGAGAAAGAG 50


RESULT  41
BI548062/c
LOCUS       BI548062                803 bp    mRNA    linear   EST 05-SEP-2001
DEFINITION  603196572F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5276392 5',
            mRNA sequence.
ACCESSION   BI548062
VERSION     BI548062.1  GI:15435361
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 803)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11697  row: o  column: 17
            High quality sequence stop: 762.
FEATURES             Location/Qualifiers
     source          1. .803
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5276392"
                     /clone_lib="NIH_MGC_95"
                     /tissue_type="hippocampus"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.5 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for  full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health).  Note: this is a NIH_MGC Library."
BASE COUNT      231 a    162 c    196 g    214 t
ORIGIN


  Query Match             8.6%;  Score 19;  DB 10;  Length 803;
  Best Local Similarity   100.0%;  Pred. No. 48;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    178 caaggggtaggagaaagag 196
          ||||||||||||||||||
Db     91 CAAGGGGTAGGAGAAAGAG 73


RESULT  42
AL542978/c
LOCUS       AL542978                813 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  AL542978 LTI_FL002_PL1 Homo sapiens cDNA clone CS0DE013YB20 5 prime
            , mRNA sequence.
ACCESSION   AL542978
VERSION     AL542978.1  GI:12875456
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 813)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .813
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DE013YB20"
                     /clone_lib="LTI_FL002_PL1"
                     /lab_host="DH10B"
                     /note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand
                     cDNA was primed with a NotI-oligo(dT) primer. Five prime
                     end enriched, double-stranded cDNA was digested with Not I
                     and cloned into the Not I and Eco RV sites of the
                     pCMVSPORT 6 vector.  Library was constructed by Life
                     Technologies. Contact : Feng Liang Life Technologies, a
                     division of Invitrogen 9800 Medical Center Drive Rockville
                     , Maryland 20850, USA Fax : (1) 301 610 8371 Email :
                     fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      231 a    156 c    192 g    232 t      2 others
```

```
ORIGIN


  Query Match             8.6%;  Score 19;  DB 9;  Length 813;
  Best Local Similarity   100.0%;  Pred. No. 48;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    178 caaggggtaggagaaagag 196
          ||||||||||||||||||||
Db     62 CAAGGGGTAGGAGAAAGAG 44


RESULT  43
BG402101/c
LOCUS       BG402101                824 bp    mRNA    linear   EST 12-MAR-2001
DEFINITION  602465634F1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4593728 5',
            mRNA sequence.
ACCESSION   BG402101
VERSION     BG402101.1  GI:13295549
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 824)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1334  row: c  column: 09
            High quality sequence stop: 569.
FEATURES             Location/Qualifiers
     source          1. .824
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4593728"
                     /clone_lib="NIH_MGC_75"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).  Average insert size 1.65
                     kb (range 0.5-4.0 kb). 15/15 colonies contained  inserts
                     by PCR. This library was enriched for full-length clones
                     and was constructed by Clontech  Laboratories (Palo Alto,
                     CA). Note: this is a NIH_MGC Library."
BASE COUNT      214 a    186 c    223 g    201 t
ORIGIN


  Query Match             8.6%;  Score 19;  DB 10;  Length 824;
  Best Local Similarity   100.0%;  Pred. No. 48;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    178 caaggggtaggagaaagag 196
          ||||||||||||||||||||
Db    100 CAAGGGGTAGGAGAAAGAG 82


RESULT  44
BI600654/c
LOCUS       BI600654                848 bp    mRNA    linear   EST 07-SEP-2001
DEFINITION  603247539F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5299400 5',
            mRNA sequence.
ACCESSION   BI600654
VERSION     BI600654.1  GI:15493593
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 848)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11757  row: n  column: 09
            High quality sequence stop: 686.
FEATURES             Location/Qualifiers
     source          1. .848
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5299400"
                     /clone_lib="NIH_MGC_96"
                     /tissue_type="hypothalamus"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.3 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length  clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      216 a    199 c    205 g    228 t
ORIGIN


  Query Match             8.6%;  Score 19;  DB 10;  Length 848;
  Best Local Similarity   100.0%;  Pred. No. 48;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    178 caaggggtaggagaaagag 196
          ||||||||||||||||||||
Db     68 CAAGGGGTAGGAGAAAGAG 50


RESULT  45
BI764049/c
LOCUS       BI764049                879 bp    mRNA    linear   EST 25-SEP-2001
DEFINITION  603043255F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5183797 5',
            mRNA sequence.
ACCESSION   BI764049
VERSION     BI764049.1  GI:15755627
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 879)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
```

```
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM11458  row: m  column: 14
           High quality sequence stop: 829.
FEATURES             Location/Qualifiers
     source          1. .879
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5183797"
                     /clone_lib="NIH_MGC_116"
                     /lab_host="DH10B"
                     /note="Organ: pooled colon, kidney, stomach; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 3 colons, age 26 yo male, 49 yo
                     female, 71 yo male colon; 46 yo male kidney, and pool of 2
                     stomachs, 62 yo male and 70 yo female.  Library is
                     oligo-dT primed and  directionally cloned (EcoRV site is
                     destroyed upon cloning).  Average insert size 1.4 kb,
                     insert size range 1-3 kb.  Library is normalized and
                     enriched for full-length clones and was constructed by C.
                     Gruber  (Invitrogen).  Research Genetics tracking code
                     023. Note: this is a NIH_MGC Library."
BASE COUNT      254 a    175 c    198 g    252 t
ORIGIN


  Query Match             8.6%;  Score 19;  DB 10;  Length 879;
  Best Local Similarity   100.0%;  Pred. No. 48;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    178 caaggggtaggagaaagag 196
          ||||||||||||||||||
Db     45 CAAGGGGTAGGAGAAAGAG 27


Search completed: September 20, 2002, 04:07:29
Job time: 13783 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 06:32:15 ; Search time 5250.46 Seconds
(without alignments)
633.720 Million cell updates/sec

Title: US-09-846-456-5
Perfect score: 159
Sequence: 1 ttaatgaccagccacgggcg.....ctttcagaagaagacaaaca 159

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	159	100.0	159	6	AX351033	AX351033 Sequence
2	159	100.0	357	6	AX351030	AX351030 Sequence
3	77	48.4	10442	6	AX060713	AX060713 Sequence
4	77	48.4	10442	6	AX060892	AX060892 Sequence
5	77	48.4	10442	9	AF285167	AF285167 Homo sapi
6	77	48.4	10474	6	AX060719	AX060719 Sequence
7	77	48.4	10474	6	AX060721	AX060721 Sequence
8	77	48.4	10474	6	AX060898	AX060898 Sequence
9	77	48.4	10474	6	AX060900	AX060900 Sequence
10	77	48.4	149034	9	AF275948	AF275948 Homo sapi
11	60	37.7	200	9	AF258623S2	AF258624 Homo sapi
12	60	37.7	298	9	AB037924	AB037924 Homo sapi
13	60	37.7	446	6	AX127764	AX127764 Sequence
14	60	37.7	446	6	AX139751	AX139751 Sequence
15	60	37.7	480	9	HSA252277	AJ252277 Homo sapi
16	60	37.7	697	9	AF258627	AF258627 Homo sapi
17	60	37.7	6786	9	AB055982	AB055982 Homo sapi
18	60	37.7	7260	6	AX253452	AX253452 Sequence
19	60	37.7	7860	6	AX092594	AX092594 Sequence
20	60	37.7	7862	6	AX135712	AX135712 Sequence
21	60	37.7	9741	6	AX127830	AX127830 Sequence
22	60	37.7	9741	6	AX139817	AX139817 Sequence
23	60	37.7	9741	6	AX351038	AX351038 Sequence
24	60	37.7	9854	6	AX127831	AX127831 Sequence
25	60	37.7	9854	6	AX139818	AX139818 Sequence
c 26	60	37.7	129608	9	AL353685	AL353685 Human DNA
c 27	60	37.7	175064	2	AC012230	AC012230 Homo sapi
28	60	37.7	183999	6	AX092589	AX092589 Sequence
29	60	37.7	201144	9	AF287262	AF287262 Homo sapi
30	51	32.1	1556	9	AK024328	AK024328 Homo sapi
c 31	44	27.7	90698	2	AC021345	AC021345 Homo sapi
32	35	22.0	37	6	AX092843	AX092843 Sequence
33	21	13.2	21	6	AX092705	AX092705 Sequence
34	21	13.2	21	6	AX092707	AX092707 Sequence
c 35	20	12.6	151961	9	AC007388	AC007388 Homo sapi
36	19	11.9	624	6	AR066487	AR066487 Sequence
37	19	11.9	624	6	AR074100	AR074100 Sequence
38	19	11.9	624	6	AR143612	AR143612 Sequence
39	19	11.9	624	6	BD005648	BD005648 Materials
40	19	11.9	684	6	AR074136	AR074136 Sequence
41	19	11.9	684	6	BD005684	BD005684 Materials
42	19	11.9	6542	9	AB014524	AB014524 Homo sapi
c 43	19	11.9	44504	9	AC022089	AC022089 Homo sapi
c 44	19	11.9	64041	2	AC013605	AC013605 Homo sapi
45	19	11.9	66477	9	AL591133	AL591133 Human DNA

ALIGNMENTS

RESULT 1
AX351033
LOCUS AX351033 159 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 5 from Patent WO0183746.
ACCESSION AX351033
VERSION AX351033.1 GI:18616389
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Rosier-Montus,M.F., Prades,C., Lemoine,C., Naudin,L., Denefle,P., Brewer,B., Duverger,N., Remaley,A. and Santamarina-Fojo,S.
TITLE Regulatory nucleic acid sequences of the abc1 gene
JOURNAL Patent: WO 0183746-A 5 08-NOV-2001;
Aventis Pharma S.A. (FR)
FEATURES Location/Qualifiers
source 1. .159
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 32 a 47 c 48 g 32 t
ORIGIN

```
  Query Match            100.0%;  Score 159;  DB 6;  Length 159;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-78;
  Matches  159;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ttaatgaccagccacgggcgtccctgctgtcagctctggccgctgccttccagggctccc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTAATGACCAGCCACGGGCGTCCCTGCTGTCAGCTCTGGCCGCTGCCTTCCAGGGCTCCC 60

Qy     61 gagccacacgctgggcgtgctggctgagggaacatggcatgttggcctcagctgaggttg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GAGCCACACGCTGGGCGTGCTGGCTGAGGGAACATGGCATGTTGGCCTCAGCTGAGGTTG 120

Qy    121 ctgctgtggaagaacctcactttcagaagaagacaaaca 159
          |||||||||||||||||||||||||||||||||||||||
Db    121 CTGCTGTGGAAGAACCTCACTTTCAGAAGAAGACAAACA 159


RESULT   2
AX351030
LOCUS       AX351030                357 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 2 from Patent WO0183746.
ACCESSION   AX351030
VERSION     AX351030.1  GI:18616386
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Rosier-Montus,M.F., Prades,C., Lemoine,C., Naudin,L., Denefle,P.,
            Brewer,B., Duverger,N., Remaley,A. and Santamarina-Fojo,S.
  TITLE     Regulatory nucleic acid sequences of the abc1 gene
  JOURNAL   Patent: WO 0183746-A 2 08-NOV-2001;
            Aventis Pharma S.A. (FR)
FEATURES             Location/Qualifiers
     source          1. .357
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       64 a     78 c    116 g     99 t
ORIGIN


  Query Match            100.0%;  Score 159;  DB 6;  Length 357;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-78;
  Matches  159;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ttaatgaccagccacgggcgtccctgctgtcagctctggccgctgccttccagggctccc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100 TTAATGACCAGCCACGGGCGTCCCTGCTGTCAGCTCTGGCCGCTGCCTTCCAGGGCTCCC 159

Qy     61 gagccacacgctgggcgtgctggctgagggaacatggcatgttggcctcagctgaggttg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    160 GAGCCACACGCTGGGCGTGCTGGCTGAGGGAACATGGCATGTTGGCCTCAGCTGAGGTTG 219

Qy    121 ctgctgtggaagaacctcactttcagaagaagacaaaca 159
          |||||||||||||||||||||||||||||||||||||||
Db    220 CTGCTGTGGAAGAACCTCACTTTCAGAAGAAGACAAACA 258


RESULT   3
AX060713
LOCUS       AX060713                10442 bp    DNA     linear   PAT 22-JAN-2001
DEFINITION  Sequence 1 from Patent WO0078972.
ACCESSION   AX060713
VERSION     AX060713.1  GI:12406103
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 10442)
  AUTHORS   Lawn,R.M., Wade,D. and Garvin,M.
  TITLE     Regulation with binding cassette transporter protein abc1
  JOURNAL   Patent: WO 0078972-A 1 28-DEC-2000;
            CV THERAPEUTICS, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .10442
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     2898 a    2297 c    2408 g    2835 t      4 others
ORIGIN


  Query Match            48.4%;  Score 77;  DB 6;  Length 10442;
  Best Local Similarity   99.2%;  Pred. No. 2.2e-32;
  Matches  127;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     32 agctctggccgctgccttccagggctcccgagccacacgctgggcgtgctggctgaggga 91
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    229 AGCTCTGGCCGCTGCCTTCCAGGGCTCCCGAGCCACACGCTGGGCGTGCTGGCTGAGGGA 288

Qy     92 acatggcatgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaa 151
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    289 ACATGGCTTGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAA 348

Qy    152 gacaaaca 159
          ||||||||
Db    349 GACAAACA 356


RESULT   4
AX060892
LOCUS       AX060892                10442 bp    DNA     linear   PAT 22-JAN-2001
DEFINITION  Sequence 1 from Patent WO0078971.
ACCESSION   AX060892
VERSION     AX060892.1  GI:12406270
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 10442)
  AUTHORS   Lawn,R.M., Wade,D., Oram,J.F. and Garvin,M.
  TITLE     Atp binding cassette transporter protein abc1 polypeptides
  JOURNAL   Patent: WO 0078971-A 1 28-DEC-2000;
            CV THERAPEUTICS, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .10442
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     2898 a    2297 c    2408 g    2835 t      4 others
ORIGIN


  Query Match            48.4%;  Score 77;  DB 6;  Length 10442;
  Best Local Similarity   99.2%;  Pred. No. 2.2e-32;
  Matches  127;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     32 agctctggccgctgccttccagggctcccgagccacacgctgggcgtgctggctgaggga 91
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    229 AGCTCTGGCCGCTGCCTTCCAGGGCTCCCGAGCCACACGCTGGGCGTGCTGGCTGAGGGA 288

Qy     92 acatggcatgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaa 151
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    289 ACATGGCTTGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAA 348

Qy    152 gacaaaca 159
          ||||||||
Db    349 GACAAACA 356


RESULT   5
AF285167
LOCUS       AF285167                10442 bp    mRNA    linear   PRI 09-AUG-2000
```

```
DEFINITION  Homo sapiens ATP-binding cassette transporter 1 (ABCA1) mRNA,
            complete cds.
ACCESSION   AF285167
VERSION     AF285167.1  GI:9755158
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 10442)
  AUTHORS   Schwartz,K., Lawn,R.M. and Wade,D.P.
  TITLE     ABCA1 gene expression and apoA-I-mediated cholesterol efflux are
            regulated by LXR
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 10442)
  AUTHORS   Lawn,R.M., Wade,D.P., Garvin,M.R., Wang,X., Schwartz,K.,
            Porter,J.G., Seilhamer,J.J., Vaughan,A.M. and Oram,J.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUL-2000) Discovery Research, CV Therapeutics Inc.,
            3172 Porter Drive, Palo Alto, CA 94304, USA
FEATURES             Location/Qualifiers
     source          1. .10442
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="9"
                     /map="9q31"
                     /cell_type="fibroblast"
                     /tissue_type="skin"
     gene            1. .10442
                     /gene="ABCA1"
     CDS             291. .7076
                     /gene="ABCA1"
                     /codon_start=1
                     /product="ATP-binding cassette transporter 1"
                     /protein_id="AAF98175.1"
                     /db_xref="GI:9755159"
                     /translation="MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRL
                     SYPPYEQHECHFPNKAMPSAGTLPWVQGIICNANNPCFRYPTPGEAPGVVGNFNKSIV
                     ARLFSDARRLLLYSQKDTSMKDMRKVLRTLQQIKKSSSNLKLQDFLVDNETFSGFLYH
                     NLSLPKSTVDKMLRADVILHKVFLQGYQLHLTSLCNGSKSEEMIQLGDQEVSELCGLP
                     KEKLAAAERVLRSNMDILKPILRTLNSTSPFPSKELAEATKTLLHSLGTLAQELFSMR
                     SWSDMRQEVMFLTNVNSSSSSTQIYQAVSRIVCGHPEGGGLKIKSLNWYEDNNYKALF
                     GGNGTEEDAETFYDNSTTPYCNDLMKNLESSPLSRIIWKALKPLLVGKILYTPDTPAT
                     RQVMAEVNKTFQELAVFHDLEGMWEELSPKIWTFMENSQEMDLVRMLLDSRDNDHFWE
                     QQLDGLDWTAQDIVAFLAKHPEDVQSSNGSVYTWREAFNETNQAIRTISRFMECVNLN
                     KLEPIATEVWLINKSMELLDERKFWAGIVFTGITPGSIELPHHVKYKIRMDIDNVERT
                     NKIKDGYWDPGPRADPFEDMRYVWGGFAYLQDVVEQAIIRVLTGTEKKTGVYMQQMPY
                     PCYVDDIFLRVMSRSMPLFMTLAWIYSVAVIIKGIVYEKEARLKETMRIMGLDNSILW
                     FSWFISSLIPLLVSAGLLVVILKLGNLLPYSDPSVVFVFLSVFAVVTILQCFLISTLF
                     SRANLAAACGGIIYFTLYLPYVLCVAWQDYVGFTLKIFASLLSPVAFGFGCEYFALFE
                     EQGIGVQWDNLFESPVEEDGFNLTTSISMMLFDTFLYGVMTWYIEAVFPGQYGIPRPW
                     YFPCTKSYWFGEESDEKSHPGSNQKRMSEICMEEEPTHLKLGVSIQNLVKVYRDGMKV
                     AVDGLALNFYEGQITSFLGHNGAGKTTTMSILTGLFPPTSGTAYILGKDIRSEMSTIR
                     QNLGVCPQHNVLFDMLTVEEHIWFYARLKGLSEKHVKAEMEQMALDVGLPSSKLKSKT
                     SQLSGGMQRKLSVALAFVGGSKVVILDEPTAGVDPYSRRGIWELLLKYRQGRTIILST
                     HHMDEADVLGDRIAIISHGKLCCVGSSLFLKNQLGTGYYLTLVKKDVESSLSSCRNSS
                     STVSYLKKEDSVSQSSSDAGLGSDHESDTLTIDVSAISNLIRKHVSEARLVEDIGHEL
                     TYVLPYEAAKEGAFVELFHEIDDRLSDLGISSYGISETTLEEIFLKVAEESGVDAETS
                     DGTLPARRNRRAFGDKQSCLRPFTEDDAADPNDSDIDPESRETDLLSGMDGKGSYQVK
                     GWKLTQQQFVALLWKRLLIARRSRKGFFAQIVLPAVFVCIALVFSLIVPPFGKYPSLE
                     LQPWMYNEQYTFVSNDAPEDTGTLELLNALTKDPGFGTRCMEGNPIPDTPCQAGEEEW
                     TTAPVPQTIMDLFQNGNWTMQNPSPACQCSSDKIKKMLPVCPPGAGGLPPPQRKQNTA
                     DILQDLTGRNISDYLVKTYVQIIAKSLKNKIWVNEFRYGGFSLGVSNTQALPPSQEVN
                     DAIKQMKKHLKLAKDSSADRFLNSLGRFMTGLDTRNNVKVWFNNKGWHAISSFLNVIN
                     NAILRANLQKGENPSHYGITAFNHPLNLTKQQLSEVALMTTSVDVLVSICVIFAMSFV
                     PASFVVFLIQERVSKAKHLQFISGVKPVIYWLSNFVWDMCNYVVPATLVIIIFICFQQ
                     KSYVSSTNLPVLALLLLLYGWSITPLMYPASFVFKIPSTAYVVLTSVNLFIGINGSVA
                     TFVLELFTDNKLNNINDILKSVFLIFPHFCLGRGLIDMVKNQAMADALERFGENRFVS
                     PLSWDLVGRNLFAMAVEGVVFFLITVLIQYRFFIRPRPVNAKLSPLNDEDEDVRRERQ
                     RILDGGGQNDILEIKELTKIYRRKRKPAVDRICVGIPPGECFGLLGVNGAGKSSTFKM
                     LTGDTTVTRGDAFLNKNSILSNIHEVHQNMGYCPQFDAITELLTGREHVEFFALLRGV
                     PEKEVGKVGEWAIRKLGLVKYGEKYAGNYSGGNKRKLSTAMALIGGPPVVFLDEPTTG
                     MDPKARRFLWNCALSVVKEGRSVVLTSHSMEECEALCTRMAIMVNGRFRCLGSVQHLK
                     NRFGDGYTIVVRIAGSNPDLKPVQDFFGLAFPGSVLKEKHRNMLQYQLPSSLSSLARI
                     FSILSQSKKRLHIEDYSVSQTTLDQVFVNFAKDQSDDDHLKDLSLHKNQTVVDVAVLT
                     SFLQDEKVKESYV"
BASE COUNT     2898 a   2297 c   2408 g   2835 t      4 others
ORIGIN


  Query Match            48.4%;  Score 77;  DB 9;  Length 10442;
  Best Local Similarity  99.2%;  Pred. No. 2.2e-32;
  Matches  127;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     32 agctctggccgctgccttccagggctcccgagccacacgctgggcgtgctggctgaggga 91
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    229 AGCTCTGGCCGCTGCCTTCCAGGGCTCCCGAGCCACACGCTGGGCGTGCTGGCTGAGGGA 288

Qy     92 acatggcatgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaa 151
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    289 ACATGGCTTGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAA 348

Qy    152 gacaaaca 159
          ||||||||
Db    349 GACAAACA 356


RESULT   6
AX060719
LOCUS       AX060719               10474 bp    DNA     linear   PAT 22-JAN-2001
DEFINITION  Sequence 7 from Patent WO0078972.
ACCESSION   AX060719
VERSION     AX060719.1  GI:12406108
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 10474)
  AUTHORS   Lawn,R.M., Wade,D. and Garvin,M.
  TITLE     Regulation with binding cassette transporter protein abc1
  JOURNAL   Patent: WO 0078972-A 7 28-DEC-2000;
            CV THERAPEUTICS, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .10474
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     2906 a   2305 c   2416 g   2843 t      4 others
ORIGIN


  Query Match            48.4%;  Score 77;  DB 6;  Length 10474;
  Best Local Similarity  99.2%;  Pred. No. 2.2e-32;
  Matches  127;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     32 agctctggccgctgccttccagggctcccgagccacacgctgggcgtgctggctgaggga 91
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    261 AGCTCTGGCCGCTGCCTTCCAGGGCTCCCGAGCCACACGCTGGGCGTGCTGGCTGAGGGA 320

Qy     92 acatggcatgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaa 151
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    321 ACATGGCTTGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAA 380

Qy    152 gacaaaca 159
          ||||||||
Db    381 GACAAACA 388


RESULT   7
AX060721
LOCUS       AX060721               10474 bp    DNA     linear   PAT 22-JAN-2001
DEFINITION  Sequence 9 from Patent WO0078972.
ACCESSION   AX060721
VERSION     AX060721.1  GI:12406109
KEYWORDS    .
SOURCE      human.
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 10474)
  AUTHORS   Lawn,R.M., Wade,D. and Garvin,M.
  TITLE     Regulation with binding cassette transporter protein abc1
  JOURNAL   Patent: WO 0078972-A 9 28-DEC-2000;
            CV THERAPEUTICS, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .10474
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     2907 a   2304 c   2415 g   2844 t      4 others
ORIGIN


  Query Match             48.4%;  Score 77;  DB 6;  Length 10474;
  Best Local Similarity   99.2%;  Pred. No. 2.2e-32;
  Matches  127;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      32 agctctggccgctgccttccagggctcccgagccacacgctgggcgtgctggctgaggga 91
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     261 AGCTCTGGCCGCTGCCTTCCAGGGCTCCCGAGCCACACGCTGGGCGTGCTGGCTGAGGGA 320

Qy      92 acatggcatgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaa 151
           ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     321 ACATGGCTTGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAA 380

Qy     152 gacaaaca 159
           ||||||||
Db     381 GACAAACA 388


RESULT   8
AX060898
LOCUS       AX060898              10474 bp    DNA     linear   PAT 22-JAN-2001
DEFINITION  Sequence 7 from Patent WO0078971.
ACCESSION   AX060898
VERSION     AX060898.1  GI:12406275
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 10474)
  AUTHORS   Lawn,R.M., Wade,D., Oram,J.F. and Garvin,M.
  TITLE     Atp binding cassette transporter protein abc1 polypeptides
  JOURNAL   Patent: WO 0078971-A 7 28-DEC-2000;
            CV THERAPEUTICS, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .10474
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     2906 a   2305 c   2416 g   2843 t      4 others
ORIGIN


  Query Match             48.4%;  Score 77;  DB 6;  Length 10474;
  Best Local Similarity   99.2%;  Pred. No. 2.2e-32;
  Matches  127;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      32 agctctggccgctgccttccagggctcccgagccacacgctgggcgtgctggctgaggga 91
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     261 AGCTCTGGCCGCTGCCTTCCAGGGCTCCCGAGCCACACGCTGGGCGTGCTGGCTGAGGGA 320

Qy      92 acatggcatgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaa 151
           ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     321 ACATGGCTTGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAA 380

Qy     152 gacaaaca 159
           ||||||||
Db     381 GACAAACA 388


RESULT   9
AX060900
LOCUS       AX060900              10474 bp    DNA     linear   PAT 22-JAN-2001
DEFINITION  Sequence 9 from Patent WO0078971.
ACCESSION   AX060900
VERSION     AX060900.1  GI:12406276
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 10474)
  AUTHORS   Lawn,R.M., Wade,D., Oram,J.F. and Garvin,M.
  TITLE     Atp binding cassette transporter protein abc1 polypeptides
  JOURNAL   Patent: WO 0078971-A 9 28-DEC-2000;
            CV THERAPEUTICS, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .10474
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     2907 a   2304 c   2415 g   2844 t      4 others
ORIGIN


  Query Match             48.4%;  Score 77;  DB 6;  Length 10474;
  Best Local Similarity   99.2%;  Pred. No. 2.2e-32;
  Matches  127;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      32 agctctggccgctgccttccagggctcccgagccacacgctgggcgtgctggctgaggga 91
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     261 AGCTCTGGCCGCTGCCTTCCAGGGCTCCCGAGCCACACGCTGGGCGTGCTGGCTGAGGGA 320

Qy      92 acatggcatgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaa 151
           ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     321 ACATGGCTTGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAA 380

Qy     152 gacaaaca 159
           ||||||||
Db     381 GACAAACA 388


RESULT  10
AF275948
LOCUS       AF275948              149034 bp    DNA     linear   PRI 17-JUL-2000
DEFINITION  Homo sapiens ABCA1 (ABCA1) gene, complete cds.
ACCESSION   AF275948
VERSION     AF275948.1  GI:9247085
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 149034)
  AUTHORS   Santamarina-Fojo,S., Peterson,K., Knapper,C., Qiu,Y., Freeman,L.,
            Cheng,J.F., Osorio,J., Remaley,A., Yang,X.P., Haudenschild,C.,
            Prades,C., Chimini,G., Blackmon,E., Francois,T., Duverger,N.,
            Rubin,E.M., Rosier,M., Denefle,P., Fredrickson,D.S. and Brewer,H.B.
            Jr.
  TITLE     Complete genomic sequence of the human ABCA1 gene: analysis of the
            human and mouse ATP-binding cassette A promoter
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (14), 7987-7992 (2000)
  MEDLINE   20345099
REFERENCE   2  (bases 1 to 149034)
  AUTHORS   Santamarina-Fojo,S., Peterson,K.M., Knapper,C.L., Freeman,L.A.,
            Remaley,A.T., Yang,X.-P., Haudenschild,C.C., Blackmon,E.E.,
            Francois,T.L. and Brewer,H.B. Jr.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JUN-2000) Molecular Disease Branch, National
            Institutes of Heath, National Heart, Lung and Blood Institute,
            Bethesda, MD 20892, USA
FEATURES             Location/Qualifiers
     source          1. .149034
```

```
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     mRNA            join(1454. .1674,25831. .25989,40385. .40478,45012. .45153,
                     46423. .46541,67729. .67850,70831. .71007,83960. .84052,
                     89010. .89250,91962. .92101,92433. .92549,96758. .96955,
                     97702. .97907,98428. .98604,100391. .100613,102360. .102581,
                     103642. .103846,104951. .105064,106862. .107033,
                     108023. .108154,109460. .109602,109806. .109943,
                     110646. .110866,112124. .112196,113183. .113385,
                     115053. .115101,115297. .115410,116806. .116954,
                     118604. .118728,119964. .120062,123094. .123283,
                     124804. .124898,126208. .126240,127363. .127468,
                     128943. .129017,129539. .129708,130939. .131116,
                     133114. .133229,133341. .133485,134526. .134649,
                     135736. .135865,136130. .136250,137038. .137100,
                     138008. .138114,140469. .140610,140982. .141116,
                     142060. .142163,142646. .142738,143397. .143640,
                     144581. .148034)
                     /gene="ABCA1"
                     /product="ABCA1"
     5'UTR           join(1454. .1674,25831. .25923)
                     /gene="ABCA1"
     gene            1454. .148034
                     /gene="ABCA1"
     repeat_region   2406. .2678
                     /rpt_family="Alu"
     repeat_region   4240. .4509
                     /rpt_family="Alu"
     repeat_region   4998. .5270
                     /rpt_family="Alu"
     repeat_region   5563. .5839
                     /rpt_family="Alu"
     repeat_region   6615. .6877
                     /rpt_family="Alu"
     repeat_region   8800. .9046
                     /rpt_family="Alu"
     repeat_region   9998. .10279
                     /rpt_family="Alu"
     repeat_region   11859. .12133
                     /rpt_family="Alu"
     repeat_region   12810. .12902
                     /rpt_family="Alu"
     repeat_region   15220. .15399
                     /rpt_family="Alu"
     repeat_region   15508. .15787
                     /rpt_family="Alu"
     repeat_region   18600. .18860
                     /rpt_family="Alu"
     repeat_region   20950. .21206
                     /rpt_family="Alu"
     repeat_region   25038. .25314
                     /rpt_family="Alu"
     CDS             join(25924. .25989,40385. .40478,45012. .45153,46423. .46541,
                     67729. .67850,70831. .71007,83960. .84052,89010. .89250,
                     91962. .92101,92433. .92549,96758. .96955,97702. .97907,
                     98428. .98604,100391. .100613,102360. .102581,103642. .103846,
                     104951. .105064,106862. .107033,108023. .108154,
                     109460. .109602,109806. .109943,110646. .110866,
                     112124. .112196,113183. .113385,115053. .115101,
                     115297. .115410,116806. .116954,118604. .118728,
                     119964. .120062,123094. .123283,124804. .124898,
                     126208. .126240,127363. .127468,128943. .129017,
                     129539. .129708,130939. .131116,133114. .133229,
                     133341. .133485,134526. .134649,135736. .135865,
                     136130. .136250,137038. .137100,138008. .138114,
                     140469. .140610,140982. .141116,142060. .142163,
                     142646. .142738,143397. .143640,144581. .144721)
                     /gene="ABCA1"
                     /codon_start=1
                     /product="ABCA1"
                     /protein_id="AAF86276.1"
                     /db_xref="GI:9247086"
                     /translation="MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRL
                     SYPPYEQHECHFPNKAMPSAGTLPWVQGIICNANNPCFRYPTPGEAPGVVGNFNKSIV
                     ARLFSDARRLLLYSQKDTSMKDMRKVLRTLQQIKKSSSNLKLQDFLVDNETFSGFLYH
                     NLSLPKSTVDKMLRADVILHKVFLQGYQLHLTSLCNGSKSEEMIQLGDQEVSELCGLP
                     REKLAAAERVLRSNMDILKPILRTLNSTSPFPSKELAEATKTLLHSLGTLAQELFSMR
                     SWSDMRQEVMFLTNVNSSSSSTQIYQAVSRIVCGHPEGGGLKIKSLNWYEDNNYKALF
                     GGNGTEEDAETFYDNSTTPYCNDLMKNLESSPLSRIIWKALKPLLVGKILYTPDTPAT
                     RQVMAEVNKTFQELAVFHDLEGMWEELSPKIWTFMENSQEMDLVRMLLDSRDNDHFWE
                     QQLDGLDWTAQDIVAFLAKHPEDVQSSNGSVYTWREAFNETNQAIRTISRFMECVNLN
                     KLEPIATEVWLINKSMELLDERKFWAGIVFTGITPGSIELPHHVKYKIRMDIDNVERT
                     NKIKDGYWDPGPRADPFEDMRYVWGGFAYLQDVVEQAIIRVLTGTEKKTGVYMQQMPY
                     PCYVDDIFLRVMSRSMPLFMTLAWIYSVAVIIKGIVYEKEARLKETMRIMGLDNSILW
                     FSWFISSLIPLLVSAGLLVVILKLGNLLPYSDPSVVFVFLSVFAVVTILQCFLISTLF
                     SRANLAAACGGIIYFTLYLPYVLCVAWQDYVGFTLKIFAXLLSPVAFGFGCEYFALFE
                     EQGIGVQWDNLFESPVEEDGFNLTTSVSMMLFDTFLYGVMTWYIEAVFPGQYGIPRPW
                     YFPCTKSYWFGEESDEKSHPGSNQKRISEICMEEEPTHLKLGVSIQNLVKVYRDGMKV
                     AVDGLALNFYEGQITSFLGHNGAGKTTTMSILTGLFPPTSGTAYILGKDIRSEMSTIR
                     QNLGVCPQHNVLFDMLTVEEHIWFYARLKGLSEKHVKAEMEQMALDVGLPSSKLKSKT
                     SQLSGGMQRKLSVALAFVGGSKVVILDEPTAGVDPYSRRGIWELLLKYRQGRTIILST
                     HHMDEADVLGDRIAIISHGKLCCVGSSLFLKNQLGTGYYLTLVKKDVESSLSSCRNSS
                     STVSYLKKEDSVSQSSSDAGLGSDHESDTLTIDVSAISNLIRKHVSEARLVEDIGHEL
                     TYVLPYEAAKEGAFVELFHEIDDRLSDLGISSYGISETTLEEIFLKVAEESGVDAETS
                     DGTLPARRNRRAFGDKQSCLRPFTEDDAADPNDSDIDPESRETDLLSGMDGKGSYQVK
                     GWKLTQQQFVALLWKRLLIARRSRKGFFAQIVLPAVFVCIALVFSLIVPPFGKYPSLE
                     LQPWMYNEQYTFVSNDAPEDTGTLELLNALTKDPGFGTRCMEGNPIPDTPCQAGEEEW
                     TTAPVPQTIMDLFQNGNWTMQNPSPACQCSSDKIKKMLPVCPPGAGGLPPPQRKQNTA
                     DILQDLTGRNISDYLVKTYVQIIAKSLKNKIWVNEFRYGGFSLGVSNTQALPPSQEVN
                     DAXKQMKKHLKLAKDSSADRFLNSLGRFMTGLDTRNNVKVWFNNKGWHAISSFLNVIN
                     NAILRANLQKGENPSHYGITAFNHPLNLTKQQLSEVAXMTTSVDVLVSICVIFAMSFV
                     PASFVVFLIQERVSKAKHLQFISGVKPVIYWLSNFVWDMCNYVVPATLVIIIFICFQQ
                     KSYVSSTNLPVLALLLLLYGWSITPLMYPASFVFKIPSTAYVVLTSVNLFIGINGSVA
                     TFVLELFTDNKLNNINDILKSVFLIFPHFCLGRGLIDMVKNQAMADALERFGENRFVS
                     PLSWDLVGRNLFAMAVEGVVFFLITVLIQYRFFIRPRPVNAKLSPLNDEDEDVRRERQ
                     RILDGGGQNDILEIKELTKIYRRKRKPAVDRICVGIPPGECFGLLGVNGAGKSSTFKM
                     LTGDTTVTRGDAFLNXNSILSNIHEVHQNMGYCPQFDAITELLTGREHVEFFALLRGV
                     PEKEVGKVGEWAIRKLGLVKYGEKYAGNYSGGNKRKLSTAMALIGGPPVVFLDEPTTG
                     MDPKARRFLWNCALSVVKEGRSVVLTSHSMEECEALCTRMAIMVNGRFRCLGSVQHLK
                     NRFGDGYTIVVRIAGSNPDLKPVQDFFGLAFPGSVXKEKHRNMLQYQLPSSLSSLARI
                     FSILSQSKKRLHIEDYSVSQTTLDQVFVNFAKDQSDDDHLKDLSLHKNQTVVDVAVLT
                     SFLQDEKVKESYV"
     repeat_region   29916. .30180
                     /rpt_family="Alu"
     repeat_region   31050. .31314
                     /rpt_family="Alu"
     repeat_region   31692. .31973
                     /rpt_family="Alu"
     repeat_region   36167. .36440
                     /rpt_family="Alu"
     repeat_region   37617. .37892
                     /rpt_family="Alu"
     repeat_region   37999. .38166
                     /rpt_family="Alu"
     repeat_region   38172. .38269
                     /rpt_family="Alu"
     repeat_region   38995. .39271
                     /rpt_family="Alu"
     STS             41547. .41565
                     /gene="ABCA1"
                     /note="A009X28"
     repeat_region   41593. .41835
                     /rpt_family="Alu"
     STS             41657. .41679
                     /gene="ABCA1"
                     /note="A009X28"
     repeat_region   42320. .42584
                     /rpt_family="Alu"
     repeat_region   44349. .44626
                     /rpt_family="Alu"
     repeat_region   48473. .48669
                     /rpt_family="Alu"
     repeat_region   50547. .50614
                     /rpt_family="Alu"
     repeat_region   54285. .54418
                     /rpt_family="Alu"
```

```
     repeat_region   54849. .54988
                     /rpt_family="Alu"
     repeat_region   55029. .55226
                     /rpt_family="Alu"
     repeat_region   55297. .55344
                     /rpt_family="Alu"
     repeat_region   58501. .58935
                     /note="LTR"
                     /rpt_family="HERV"
     repeat_region   61246. .61489
                     /rpt_family="Alu"
     repeat_region   62812. .63077
                     /rpt_family="Alu"
     repeat_region   64082. .64306
                     /rpt_family="Alu"
     repeat_region   67123. .67382
                     /rpt_family="Alu"
     repeat_region   68499. .68731
                     /rpt_family="Alu"
     repeat_region   69481. .69760
                     /rpt_family="Alu"
     repeat_region   71998. .72233
                     /rpt_family="Alu"
     repeat_region   72244. .72501
                     /rpt_family="Alu"

  Query Match             48.4%;  Score 77;  DB 9;  Length 149034;
  Best Local Similarity   99.2%;  Pred. No. 2e-32;
  Matches  127;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      32 agctctggccgctgccttccagggctcccgagccacacgctgggcgtgctggctgaggga 91
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25862 AGCTCTGGCCGCTGCCTTCCAGGGCTCCCGAGCCACACGCTGGGCGTGCTGGCTGAGGGA 25921

Qy      92 acatggcatgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaa 151
           ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25922 ACATGGCTTGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAA 25981

Qy     152 gacaaaca 159
           ||||||||
Db   25982 GACAAACA 25989


RESULT  11
AF258623S2
LOCUS       AF258623S2              200 bp    DNA     linear   PRI 23-JUN-2000
DEFINITION  Homo sapiens ATP binding cassette transporter 1 (ABCA1) gene, exon
            2.
ACCESSION   AF258624
VERSION     AF258624.1  GI:7769714
KEYWORDS    .
SEGMENT     2 of 4
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 200)
  AUTHORS   Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
            Aouizerat,B.E., Fielding,C.J. and Kane,J.P.
  TITLE     Analysis of hABC1 gene 5' end: additional peptide sequence,
            promoter region, and four polymorphisms
  JOURNAL   Biochem. Biophys. Res. Commun. 271 (2000) In press
REFERENCE   2  (bases 1 to 200)
  AUTHORS   Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
            Aouizerat,B.E., Fielding,C.J. and Kane,J.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2000) Cardiovascular Research Institute,
            University of California, San Francisco, 505 Parnassus Avenue, San
            Francisco, CA 94143-0130, USA
  REMARK    Sequence update by submitter
FEATURES             Location/Qualifiers
     source          1. .200
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="9"
                     /map="9q31"
     exon            22. .179
                     /gene="ABCA1"
                     /number=2
BASE COUNT       37 a     56 c     57 g     50 t
ORIGIN


  Query Match             37.7%;  Score 60;  DB 9;  Length 200;
  Best Local Similarity   100.0%;  Pred. No. 9.9e-23;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     120 TGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAAGACAAACA 179


RESULT  12
AB037924
LOCUS       AB037924                298 bp    mRNA    linear   PRI 12-OCT-2000
DEFINITION  Homo sapiens mRNA for ABC1, partial cds.
ACCESSION   AB037924
VERSION     AB037924.1  GI:9711458
KEYWORDS    ABC1.
SOURCE      Homo sapiens placenta cDNA to mRNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Zhao,L.X., Zhou,C.J., Tanaka,A., Nakata,M., Hirabayashi,T.,
            Amachi,T., Shioda,S., Ueda,K. and Inagaki,N.
  TITLE     Cloning, characterization and tissue distribution of the rat
            ATP-binding cassette (ABC) transporter ABC2/ABCA2
  JOURNAL   Biochem. J. 350 (Pt 3), 865-872 (2000)
   PUBMED   10970803
REFERENCE   2  (bases 1 to 298)
  AUTHORS   Ueda,K., Kioka,N. and Tanaka,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-FEB-2000) Kazumitsu Ueda, Kyoto University Graduate
            School of Agriculture, Division of Applied Life Sciences;
            Kitashirakawa, Sakyo-ku, Kyoto, kyoto 606-8502, Japan
            (E-mail:uedak@kais.kyoto-u.ac.jp, Tel:81-75-753-6105,
            Fax:81-75-753-6104)
FEATURES             Location/Qualifiers
     source          1. .298
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /tissue_type="placenta"
     gene            88. .298
                     /gene="ABC1"
     CDS             88. .>298
                     /gene="ABC1"
                     /codon_start=1
                     /product="ABC1"
                     /protein_id="BAB07875.1"
                     /db_xref="GI:9711459"
                     /translation="MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRL
                     SYPPYEQHECHFPNKAMPSAGTLPWV"
BASE COUNT       60 a     87 c     77 g     72 t      2 others
ORIGIN


  Query Match             37.7%;  Score 60;  DB 9;  Length 298;
  Best Local Similarity   100.0%;  Pred. No. 9.8e-23;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94 TGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAAGACAAACA 153
```

```
RESULT  13
AX127764
LOCUS       AX127764                446 bp    DNA     linear   PAT 15-MAY-2001
DEFINITION  Sequence 3 from Patent WO0130848.
ACCESSION   AX127764
VERSION     AX127764.1  GI:14134411
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 446)
  AUTHORS   Denefle,P., Rosier-Montus,M.F., Arnould-Reguigne,I., Prades,C.,
            Naudin,L., Lemoine,C., Duverger,N., Jaye,M., Searfoss,G.H.,
            Remaley,A., Brewer,H.B. and Dean,M.
  TITLE     Nucleic acids of the human abc1 gene and their therapeutic and
            diagnostic application
  JOURNAL   Patent: WO 0130848-A 3 03-MAY-2001;
            Aventis Pharma S.A. (FR)
FEATURES             Location/Qualifiers
     source          1. .446
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Oligonucleotide Primer"
BASE COUNT       96 a    123 c    112 g    115 t
ORIGIN


  Query Match           37.7%;  Score 60;  DB 6;  Length 446;
  Best Local Similarity   100.0%;  Pred. No. 9.7e-23;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     191 TGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAAGACAAACA 250


RESULT  14
AX139751
LOCUS       AX139751                446 bp    DNA     linear   PAT 30-MAY-2001
DEFINITION  Sequence 3 from Patent EP1096012.
ACCESSION   AX139751
VERSION     AX139751.1  GI:14275333
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 446)
  AUTHORS   Denefle,P., Rosier-Montus,M.F., Arnould-Reguigne,I., Prades,C.,
            Naudin,L., Lemoine,C., Duverger,N., Jaye,M., searfoss Iii,G.H.,
            Remaley,A., Brewer,H.B. and Dean,M.
  TITLE     Nucleics acids of the human abc1 gene and their therapeutic and
            diagnostic application
  JOURNAL   Patent: EP 1096012-A 3 02-MAY-2001;
            Aventis Pharma S.A. (FR)
FEATURES             Location/Qualifiers
     source          1. .446
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Oligonucleotide Primer"
BASE COUNT       96 a    123 c    112 g    115 t
ORIGIN


  Query Match           37.7%;  Score 60;  DB 6;  Length 446;
  Best Local Similarity   100.0%;  Pred. No. 9.7e-23;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     191 TGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAAGACAAACA 250


RESULT  15
HSA252277
LOCUS       HSA252277               480 bp    DNA     linear   PRI 08-JAN-2001
DEFINITION  Homo sapiens partial ABC-1 gene for ATP-binding cassette
            transporter-1, exon 2.
ACCESSION   AJ252277
VERSION     AJ252277.1  GI:12140344
KEYWORDS    ABC-1 gene; ATP-binding cassette transporter-1.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 480)
  AUTHORS   Porsch-Oezcueruemez,M., Langmann,T. and Schmitz,G.
  TITLE     Cloning and Characterization of the human ATP-binding Cassette
            Transporter-1 (ABC-1) Promoter
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 480)
  AUTHORS   Porsch-Oezcueruemez,M.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JAN-2000) Porsch-Oezcueruemez M.K., Institute for
            Clinical Chemistry, University of Regensburg,
            Franz-Josef-Strauss-Allee 11, 93042 Regensburg, GERMANY
FEATURES             Location/Qualifiers
     source          1. .480
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="9"
                     /map="9q31"
                     /cell_type="leukocyte"
     exon            189. .346
                     /gene="ABC-1"
                     /number=2
     5'UTR           189. .280
                     /gene="ABC-1"
                     /function="cholesterol efflux regulatory proteiin"
     gene            189. .346
                     /gene="ABC-1"
     CDS             281. .>346
                     /gene="ABC-1"
                     /function="cholesterol efflux regulatory protein"
                     /codon_start=1
                     /product="ATP-binding cassette transporter-1"
                     /protein_id="CAC21428.1"
                     /db_xref="GI:12140345"
                     /translation="MACWPQLRLLLWKNLTFRRRQT"
BASE COUNT       89 a    102 c    155 g    134 t
ORIGIN


  Query Match           37.7%;  Score 60;  DB 9;  Length 480;
  Best Local Similarity   100.0%;  Pred. No. 9.6e-23;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     287 TGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAAGACAAACA 346


RESULT  16
AF258627
LOCUS       AF258627                697 bp    mRNA    linear   PRI 11-MAY-2000
DEFINITION  Homo sapiens ATP binding cassette transporter 1 (ABCA1) mRNA,
            partial cds.
ACCESSION   AF258627
VERSION     AF258627.1  GI:7769707
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 697)
  AUTHORS   Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
            Aouizerat,B.E., Fielding,C.J. and Kane,J.P.
```

```
  TITLE     Analysis of hABC1 gene 5' end: additional peptide sequence,
            promoter region, and four polymorphisms
  JOURNAL   Biochem. Biophys. Res. Commun. 271 (2000) In press
REFERENCE   2  (bases 1 to 697)
  AUTHORS   Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
            Aouizerat,B.E., Fielding,C.J. and Kane,J.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-APR-2000) Cardiovascular Research Institute,
            University of California, San Francisco, 505 Parnassus Avenue, San
            Francisco, CA 94143-0130, USA
FEATURES             Location/Qualifiers
     source          1. .697
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="9"
                     /map="9q31"
                     /tissue_type="placenta"
     gene            1. .>697
                     /gene="ABCA1"
                     /note="ABC1"
     CDS             396. .>697
                     /gene="ABCA1"
                     /note="membrane-bound"
                     /codon_start=1
                     /product="ATP binding cassette transporter 1"
                     /protein_id="AAF69513.1"
                     /db_xref="GI:7769708"
                     /translation="MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRL
                     SYPPYEQHECHFPNKAMPSAGTLPWVQGIICNANNPCFRYPTPGEAPGVVGNFNKS"
BASE COUNT      152 a    198 c    190 g    156 t      1 others
ORIGIN


  Query Match             37.7%;  Score 60;  DB 9;  Length 697;
  Best Local Similarity   100.0%;  Pred. No. 9.5e-23;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    402 TGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAAGACAAACA 461


RESULT  17
AB055982
LOCUS       AB055982                6786 bp    mRNA    linear   PRI 18-AUG-2001
DEFINITION  Homo sapiens mRNA for ABCA1, complete cds.
ACCESSION   AB055982
VERSION     AB055982.1  GI:15212106
KEYWORDS    .
SOURCE      Homo sapiens cDNA to mRNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 6786)
  AUTHORS   Tanaka,A.R., Abe-Dohmae,S., Arakawa,R., Sadanami,K., Kidera,A.,
            Kioka,N., Amachi,T., Yokoyama,S. and Ueda,K.
  TITLE     A new topological model of functional human ABCA1-Signal peptide
            cleavage and glycosylation of a large extracellular domain
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 6786)
  AUTHORS   Ueda,K., Kioka,N. and Tanaka,A.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2001) Kazumitsu Ueda, Kyoto University Graduate
            School, Applied Life Sciences; Kitashirakawa, Kyoto Sakyo-ku, Kyoto
            606-8502, Japan (E-mail:uedak@kais.kyoto-u.ac.jp,
            Tel:81-75-753-6105, Fax:81-75-753-6104)
FEATURES             Location/Qualifiers
     source          1. .6786
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            1. .6786
                     /gene="ABCA1"
     CDS             1. .6786
                     /gene="ABCA1"
                     /codon_start=1
                     /product="ABCA1"
                     /protein_id="BAB63210.1"
                     /db_xref="GI:15212107"
                     /translation="MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRL
                     SYPPYEQHECHFPNKAMPSAGTLPWVQGIICNANNPCFRYPTPGEAPGVVGNFNKSIV
                     ARLFSDARRLLLYSQKDTSMKDMRKVLRTLQQIKKSSSNLKLQDFLVDNETFSGFLYH
                     NLSLPKSTVDKMLRADVILHKVFLQGYQLHLTSLCNGSKSEEMIQLGDQEVSELCGLP
                     REKLAAAERVLRSNMDILKPILRTLNSTSPFPSKELAEATKTLLHSLGTLAQELFSMR
                     SWSDMRQEVMFLTNVNSSSSSTQIYQAVSRIVCGHPEGGGLKIKSLNWYEDNNYKALF
                     GGNGTEEDAETFYDNSTTPYCNDLMKNLESSPLSRIIWKALKPLLVGKILYTPDTPAT
                     RQVMAEVNKTFQELAVFHDLEGMWEELSPKIWTFMENSQEMDLVRMLLDSRDNDHFWE
                     QQLDGLDWTAQDIVAFLAKHPEDVQSSNGSVYTWREAFNETNQAIRTISRFMECVNLN
                     KLEPIATEVWLINKSMELLDERKFWAGIVFTGITPGSIELPHHVKYKIRMDIDNVERT
                     NKIKDGYWDPGPRADPFEDMRYVWGGFAYLQDVVEQAIIRVLTGTEKKTGVYMQQMPY
                     PCYVDDIFLRVMSRSMPLFMTLAWIYSVAVIIKGIVYEKEARLKETMRIMGLDNSILW
                     FSWFISSLIPLLVSAGLLVVILKLGNLLPYSDPSVVFVFLSVFAVVTILQCFLISTLF
                     SRANLAAACGGIIYFTLYLPYVLCVAWQDYVGFTLKIFASLLSPVAFGFGCEYFALFE
                     EQGIGVQWDNLFESPVEEDGFNLTTSVSMMLFDTFLYGVMTWYIEAVFPGQYGIPRPW
                     YFPCTKSYWFGEESDEKSHPGSNQKRISEICMEEEPTHLKLGVSIQNLVKVYRDGMKV
                     AVDGLALNFYEGQITSFLGHNGAGKTTTMSILTGLFPPTSGTAYILGKDIRSEMSTIR
                     QNLGVCPQHNVLFDMLTVEEHIWFYARLKGLSEKHVKAEMEQMALDVGLPSSKLKSKT
                     SQLSGGMQRKLSVALAFVGGSKVVILDEPTAGVDPYSRRGIWELLLKYRQGRTIILST
                     HHMDEADVLGDRIAIISHGKLCCVGSSLFLKNQLGTGYYLTLVKKDVESSLSSCRNSS
                     STVSYLKKEDSVSQSSSDAGLGSDHESDTLTIDVSAISNLIRKHVSEARLVEDIGHEL
                     TYVLPYEAAKEGAFVELFHEIDDRLSDLGISSYGISETTLEEIFLKVAEESGVDAETS
                     DGTLPARRNRRAFGDKQSCLRPFTEDDAADPNDSDIDPESRETDLLSGMDGKGSYQVK
                     GWKLTQQQFVALLWKRLLIARRSRKGFFAQIVLPAVFVCIALVFSLIVPPFGKYPSLE
                     LQPWMYNEQYTFVSNDAPEDTGTLELLNALTKDPGFGTRCMEGNPIPDTPCQAGEEEW
                     TTAPVPQTIMDLFQNGNWTMQNPSPACQCSSDKIKKMLPVCPPGAGGLPPPQRKQNTA
                     DILQDLTGRNISDYLVKTYVQIIAKSLKNKIWVNEFRYGGFSLGVSNTQALPPSQEVN
                     DAIKQMKKHLKLAKDSSADRFLNSLGRFMTGLDTKNNVKVWFNNKGWHAISSFLNVIN
                     NAILRANLQKGENPSHYGITAFNHPLNLTKQQLSEVALMTTSVDVLVSICVIFAMSFV
                     PASFVVFLIQERVSKAKHLQFISGVKPVIYWLSNFVWDMCNYVVPATLVIIIFICFQQ
                     KSYVSSTNLPVLALLLLLYGWSITPLMYPASFVFKIPSTAYVVLTSVNLFIGINGSVA
                     TFVLELFTDNKLNNINDILKSVFLIFPHFCLGRGLIDMVKNQAMADALERFGENRFVS
                     PLSWDLVGRNLFAMAVEGVVFFLITVLIQYRFFIRPRPVNAKLSPLNDEDEDVRRERQ
                     RILDGGGQNDILEIKELTKIYRRKRKPAVDRICVGIPPGECFGLLGVNGAGKSSTFKM
                     LTGDTTVTRGDAFLNKNSILSNIHEVHQNMGYCPQFDAITELLTGREHVEFFALLRGV
                     PEKEVGKVGEWAIRKLGLVKYGEKYAGNYSGGNKRKLSTAMALIGGPPVVFLDEPTTG
                     MDPKARRFLWNCALSVVKEGRSVVLTSHSMEECEALCTRMAIMVNGRFRCLGSVQHLK
                     NRFGDGYTIVVRIAGSNPDLKPVQDFFGLAFPGSVLKEKHRNMLQYQLPSSLSSLARI
                     FSILSQSKKRLHIEDYSVSQTTLDQVFVNFAKDQSDDDHLKDLSLHKNQTVVDVAVLT
                     SFLQDEKVKESYV"
BASE COUNT     1724 a   1643 c   1759 g   1660 t
ORIGIN


  Query Match             37.7%;  Score 60;  DB 9;  Length 6786;
  Best Local Similarity   100.0%;  Pred. No. 8.8e-23;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7 TGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAAGACAAACA 66


RESULT  18
AX253452
LOCUS       AX253452                7260 bp    DNA     linear   PAT 10-OCT-2001
DEFINITION  Sequence 3 from Patent WO0170810.
ACCESSION   AX253452
VERSION     AX253452.1  GI:16073979
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 7260)
  AUTHORS   Schmitz,G. and Bodzioch,M.
  TITLE     Atp binding cassette transporter 1 (abc1) gene polymorphisms and
            uses thereof for the diagnosis and treatment of lipid,
```

```
          cardiovascular or inflammatory disorders
  JOURNAL   Patent: WO 0170810-A 3 27-SEP-2001;
            Bayer Aktiengesellschaft (DE)
FEATURES             Location/Qualifiers
     source          1. .7260
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1834 a   1765 c   1905 g   1756 t
ORIGIN


  Query Match             37.7%;  Score 60;  DB 6;  Length 7260;
  Best Local Similarity   100.0%;  Pred. No. 8.8e-23;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    327 TGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAAGACAAACA 386


RESULT  19
AX092594
LOCUS       AX092594                7860 bp    DNA     linear   PAT 21-MAR-2001
DEFINITION  Sequence 6 from Patent WO0115676.
ACCESSION   AX092594
VERSION     AX092594.1  GI:13444651
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 7860)
  AUTHORS   Hayden,M.R., Brooks-Wilson,A.R., Pimstone,S.N. and Clee,S.M.
  TITLE     Compositions and methods for modulating hdl cholesterol and
            triglyceride levels
  JOURNAL   Patent: WO 0115676-A 6 08-MAR-2001;
            University of British Columbia (CA) ; Xenon Genetics Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .7860
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     2014 a   1860 c   2008 g   1978 t
ORIGIN


  Query Match             37.7%;  Score 60;  DB 6;  Length 7860;
  Best Local Similarity   100.0%;  Pred. No. 8.8e-23;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81 TGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAAGACAAACA 140


RESULT  20
AX135712
LOCUS       AX135712                7862 bp    DNA     linear   PAT 29-MAY-2001
DEFINITION  Sequence 1 from Patent WO0132184.
ACCESSION   AX135712
VERSION     AX135712.1  GI:14271961
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 7862)
  AUTHORS   Attie,A.D., Cook,M., Gray-Keller,M.P., Hayden,M.R., Pimstone,S. and
            Brooks-Wilson,A.
  TITLE     Abc1 modulation for the modulation of cholesterol transport
  JOURNAL   Patent: WO 0132184-A 1 10-MAY-2001;
            WISCONSIN ALUMNI RESEARCH FOUNDATION (US)
FEATURES             Location/Qualifiers
     source          1. .7862
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     2013 a   1861 c   2010 g   1978 t
ORIGIN


  Query Match             37.7%;  Score 60;  DB 6;  Length 7862;
  Best Local Similarity   100.0%;  Pred. No. 8.8e-23;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81 TGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAAGACAAACA 140


RESULT  21
AX127830
LOCUS       AX127830                9741 bp    DNA     linear   PAT 15-MAY-2001
DEFINITION  Sequence 69 from Patent WO0130848.
ACCESSION   AX127830
VERSION     AX127830.1  GI:14134477
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 9741)
  AUTHORS   Denefle,P., Rosier-Montus,M.F., Arnould-Reguigne,I., Prades,C.,
            Naudin,L., Lemoine,C., Duverger,N., Jaye,M., Searfoss,G.H.,
            Remaley,A., Brewer,H.B. and Dean,M.
  TITLE     Nucleic acids of the human abc1 gene and their therapeutic and
            diagnostic application
  JOURNAL   Patent: WO 0130848-A 69 03-MAY-2001;
            Aventis Pharma S.A. (FR)
FEATURES             Location/Qualifiers
     source          1. .9741
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     2650 a   2180 c   2290 g   2620 t      1 others
ORIGIN


  Query Match             37.7%;  Score 60;  DB 6;  Length 9741;
  Best Local Similarity   100.0%;  Pred. No. 8.7e-23;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    191 TGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAAGACAAACA 250


RESULT  22
AX139817
LOCUS       AX139817                9741 bp    DNA     linear   PAT 30-MAY-2001
DEFINITION  Sequence 69 from Patent EP1096012.
ACCESSION   AX139817
VERSION     AX139817.1  GI:14275399
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 9741)
  AUTHORS   Denefle,P., Rosier-Montus,M.F., Arnould-Reguigne,I., Prades,C.,
            Naudin,L., Lemoine,C., Duverger,N., Jaye,M., searfoss Iii,G.H.,
            Remaley,A., Brewer,H.B. and Dean,M.
  TITLE     Nucleics acids of the human abc1 gene and their therapeutic and
            diagnostic application
  JOURNAL   Patent: EP 1096012-A 69 02-MAY-2001;
            Aventis Pharma S.A. (FR)
FEATURES             Location/Qualifiers
     source          1. .9741
                     /organism="Homo sapiens"
```

```
                     /db_xref="taxon:9606"
BASE COUNT     2650 a    2180 c    2290 g    2620 t      1 others
ORIGIN


  Query Match             37.7%;  Score 60;  DB 6;  Length 9741;
  Best Local Similarity   100.0%;  Pred. No. 8.7e-23;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     191 TGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAAGACAAACA 250


RESULT  23
AX351038
LOCUS       AX351038               9741 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 10 from Patent WO0183746.
ACCESSION   AX351038
VERSION     AX351038.1  GI:18616393
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Rosier-Montus,M.F., Prades,C., Lemoine,C., Naudin,L., Denefle,P.,
            Brewer,B., Duverger,N., Remaley,A. and Santamarina-Fojo,S.
  TITLE     Regulatory nucleic acid sequences of the abc1 gene
  JOURNAL   Patent: WO 0183746-A 10 08-NOV-2001;
            Aventis Pharma S.A. (FR)
FEATURES             Location/Qualifiers
     source          1. .9741
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     2650 a    2180 c    2290 g    2620 t      1 others
ORIGIN


  Query Match             37.7%;  Score 60;  DB 6;  Length 9741;
  Best Local Similarity   100.0%;  Pred. No. 8.7e-23;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     191 TGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAAGACAAACA 250


RESULT  24
AX127831
LOCUS       AX127831               9854 bp    DNA     linear   PAT 15-MAY-2001
DEFINITION  Sequence 70 from Patent WO0130848.
ACCESSION   AX127831
VERSION     AX127831.1  GI:14134478
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 9854)
  AUTHORS   Denefle,P., Rosier-Montus,M.F., Arnould-Reguigne,I., Prades,C.,
            Naudin,L., Lemoine,C., Duverger,N., Jaye,M., Searfoss,G.H.,
            Remaley,A., Brewer,H.B. and Dean,M.
  TITLE     Nucleic acids of the human abc1 gene and their therapeutic and
            diagnostic application
  JOURNAL   Patent: WO 0130848-A 70 03-MAY-2001;
            Aventis Pharma S.A. (FR)
FEATURES             Location/Qualifiers
     source          1. .9854
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     2665 a    2219 c    2334 g    2635 t      1 others
ORIGIN


  Query Match             37.7%;  Score 60;  DB 6;  Length 9854;
  Best Local Similarity   100.0%;  Pred. No. 8.7e-23;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     304 TGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAAGACAAACA 363


RESULT  25
AX139818
LOCUS       AX139818               9854 bp    DNA     linear   PAT 30-MAY-2001
DEFINITION  Sequence 70 from Patent EP1096012.
ACCESSION   AX139818
VERSION     AX139818.1  GI:14275400
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 9854)
  AUTHORS   Denefle,P., Rosier-Montus,M.F., Arnould-Reguigne,I., Prades,C.,
            Naudin,L., Lemoine,C., Duverger,N., Jaye,M., searfoss Iii,G.H.,
            Remaley,A., Brewer,H.B. and Dean,M.
  TITLE     Nucleics acids of the human abc1 gene and their therapeutic and
            diagnostic application
  JOURNAL   Patent: EP 1096012-A 70 02-MAY-2001;
            Aventis Pharma S.A. (FR)
FEATURES             Location/Qualifiers
     source          1. .9854
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     2665 a    2219 c    2334 g    2635 t      1 others
ORIGIN


  Query Match             37.7%;  Score 60;  DB 6;  Length 9854;
  Best Local Similarity   100.0%;  Pred. No. 8.7e-23;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     304 TGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAAGACAAACA 363


RESULT  26
AL353685/c
LOCUS       AL353685             129608 bp    DNA     linear   PRI 01-JUN-2001
DEFINITION  Human DNA sequence from clone RP11-31J20 on chromosome 9, complete
            sequence.
ACCESSION   AL353685
VERSION     AL353685.23  GI:14329534
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 129608)
  AUTHORS   Tracey,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Jun 8, 2001 this sequence version replaced gi:14272260.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
```

regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr9
RP11-31J20 is from the library RPCI-11.1 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
IMPORTANT: This sequence is not the entire insert of clone RP11-31J20 It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.
The true right end of clone RP11-31J20 is at 129608 in this sequence. The true right end of clone RP11-413C10 is at 2000 in this sequence.

FEATURES Location/Qualifiers

source 1. .129608
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-31J20"
/clone_lib="RPCI-11.1"

repeat_region 1358. .1653
/note="AluSc repeat: matches 1. .288 of consensus"

repeat_region 1828. .1877
/note="L2 repeat: matches 2649. .2698 of consensus"

repeat_region 2496. .2714
/note="MIR repeat: matches 12. .250 of consensus"

repeat_region 2777. .2896
/note="MIR repeat: matches 6. .128 of consensus"

repeat_region 3237. .3415
/note="L1ME repeat: matches 5696. .5821 of consensus"

repeat_region 6522. .6818
/note="AluSq repeat: matches 1. .295 of consensus"

repeat_region 7282. .7415
/note="L1MB8 repeat: matches 6040. .6173 of consensus"

repeat_region 8145. .8434
/note="AluSc repeat: matches 1. .298 of consensus"

repeat_region 12145. .12713
/note="L2 repeat: matches 1363. .1940 of consensus"

repeat_region 13890. .13969
/note="L2 repeat: matches 2611. .2701 of consensus"

repeat_region 15380. .15411
/note="16 copies 2 mer ac 87% conserved"

repeat_region 16105. .16144
/note="10 copies 4 mer caca 100% conserved"

repeat_region 16868. .17049
/note="MIR repeat: matches 64. .246 of consensus"

repeat_region 17941. .18229
/note="AluSg repeat: matches 1. .287 of consensus"

repeat_region 18259. .18553
/note="AluSq repeat: matches 1. .293 of consensus"

repeat_region 20310. .20616
/note="AluSq repeat: matches 1. .308 of consensus"

repeat_region 20957. .21107
/note="MIR repeat: matches 49. .212 of consensus"

repeat_region 21783. .22078
/note="AluSx repeat: matches 7. .302 of consensus"

repeat_region 22320. .22439
/note="MIR repeat: matches 10. .146 of consensus"

repeat_region 22533. .22839
/note="AluSx repeat: matches 1. .307 of consensus"

repeat_region 23427. .23945
/note="L2 repeat: matches 2137. .2750 of consensus"

repeat_region 24245. .24544
/note="AluSg repeat: matches 2. .302 of consensus"

repeat_region 24556. .24587
/note="8 copies 4 mer acac 96% conserved"

repeat_region 26504. .26561
/note="29 copies 2 mer ta 69% conserved"

repeat_region 26849. .26892
/note="11 copies 4 mer tata 81% conserved"

repeat_region 27998. .28103
/note="MIR repeat: matches 28. .145 of consensus"

repeat_region 28515. .28626
/note="MIR repeat: matches 17. .129 of consensus"

repeat_region 28986. .29213
/note="MIR repeat: matches 2. .245 of consensus"

repeat_region 30121. .30422
/note="AluSq repeat: matches 1. .302 of consensus"

repeat_region 31424. .31734
/note="AluSp repeat: matches 1. .308 of consensus"

repeat_region 31987. .32116
/note="MER5A repeat: matches 60. .188 of consensus"

repeat_region 32450. .32536
/note="MER5A repeat: matches 14. .106 of consensus"

repeat_region 33293. .33380
/note="22 copies 4 mer atgg 79% conserved"

repeat_region 34435. .34588
/note="MER5A repeat: matches 13. .189 of consensus"

repeat_region 34729. .34873
/note="MIR repeat: matches 49. .198 of consensus"

repeat_region 35802. .35951
/note="MIR repeat: matches 77. .248 of consensus"

repeat_region 37183. .37260
/note="2 copies 39 mer 92% conserved"

repeat_region 37673. .37980
/note="AluSc repeat: matches 1. .308 of consensus"

repeat_region 39674. .40243
/note="L1MD2 repeat: matches 5774. .6331 of consensus"

repeat_region 40256. .40534
/note="L2 repeat: matches 2256. .2533 of consensus"

repeat_region 41476. .41615
/note="MIR repeat: matches 30. .185 of consensus"

repeat_region 42010. .42194
/note="MIR repeat: matches 49. .234 of consensus"

repeat_region 44809. .45239
/note="LTR16A repeat: matches 6. .450 of consensus"

repeat_region 45359. .45486
/note="MIR repeat: matches 1. .139 of consensus"

repeat_region 46654. .46693
/note="10 copies 4 mer acac 97% conserved"

repeat_region 47118. .47429
/note="AluSx repeat: matches 1. .312 of consensus"

repeat_region 47915. .48083
/note="AluSq repeat: matches 145. .313 of consensus"

repeat_region 49097. .49181
/note="L2 repeat: matches 2625. .2710 of consensus"

repeat_region 49578. .49758
/note="MIR repeat: matches 9. .188 of consensus"

repeat_region 49847. .49873
/note="MIR repeat: matches 155. .182 of consensus"

repeat_region 49874. .50188
/note="AluJb repeat: matches 3. .311 of consensus"

repeat_region 50189. .50347
/note="MIR repeat: matches 5. .155 of consensus"

repeat_region 51677. .51977
/note="AluSx repeat: matches 1. .303 of consensus"

repeat_region 52025. .52123
/note="L2 repeat: matches 1655. .1753 of consensus"

repeat_region 54337. .54544
/note="MIR repeat: matches 25. .246 of consensus"

repeat_region 54754. .55032
/note="AluSp repeat: matches 1. .302 of consensus"

repeat_region 55042. .55343
/note="AluY repeat: matches 5. .298 of consensus"

```
     repeat_region   56027. .56232
                     /note="MIR repeat: matches 21. .259 of consensus"
     repeat_region   56233. .56286
                     /note="18 copies 3 mer tgt 72% conserved"
     repeat_region   56342. .56673
                     /note="L1MA9 repeat: matches 5893. .6307 of consensus"
     repeat_region   56750. .56889
                     /note="L2 repeat: matches 2565. .2710 of consensus"
     repeat_region   57406. .57567
                     /note="MER5A repeat: matches 4. .187 of consensus"
     repeat_region   57864. .58174
                     /note="AluSq repeat: matches 1. .311 of consensus"
     repeat_region   59443. .60004
                     /note="L2 repeat: matches 2088. .2723 of consensus"
     repeat_region   60201. .60355
                     /note="MIR repeat: matches 20. .183 of consensus"
     repeat_region   63991. .64159
                     /note="MIR repeat: matches 78. .243 of consensus"
     repeat_region   64590. .64653
                     /note="32 copies 2 mer ta 70% conserved"
     repeat_region   64593. .64656
                     /note="16 copies 4 mer atat 68% conserved"
     repeat_region   64981. .65042
                     /note="MER5A repeat: matches 15. .72 of consensus"
     repeat_region   65043. .65341
                     /note="AluSx repeat: matches 1. .299 of consensus"
     repeat_region   65342. .65441
                     /note="MER5A repeat: matches 72. .186 of consensus"
     repeat_region   66825. .67131
                     /note="AluSx repeat: matches 1. .306 of consensus"

  Query Match             37.7%;  Score 60;  DB 9;  Length 129608;
  Best Local Similarity   100.0%;  Pred. No. 8e-23;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  120383 TGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAAGACAAACA 120324


RESULT  27
AC012230/c
LOCUS       AC012230               175064 bp    DNA     linear   HTG 22-APR-2000
DEFINITION  Homo sapiens clone RP11-1M10, WORKING DRAFT SEQUENCE, 39 unordered
            pieces.
ACCESSION   AC012230
VERSION     AC012230.3  GI:7637254
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 175064)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-1M10
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 175064)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Apr 22, 2000 this sequence version replaced gi:6454033.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L2510
                Center clone name: 1_M_10
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 117571 bases at least Q40
                Consensus quality: 145749 bases at least Q30
                Consensus quality: 160940 bases at least Q20
                Insert size: 185000; agarose-fp
                Insert size: 171264; sum-of-contigs
                Quality coverage: 2.9 in Q20 bases; agarose-fp
                Quality coverage: 3.2 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 39 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     1003: contig of 1003 bp in length
            *    1004 1103: gap of      100 bp
            *    1104     2634: contig of 1531 bp in length
            *    2635 2734: gap of      100 bp
            *    2735     4415: contig of 1681 bp in length
            *    4416 4515: gap of      100 bp
            *    4516     5785: contig of 1270 bp in length
            *    5786 5885: gap of      100 bp
            *    5886     7879: contig of 1994 bp in length
            *    7880 7979: gap of      100 bp
            *    7980     9686: contig of 1707 bp in length
            *    9687 9786: gap of      100 bp
            *    9787    12253: contig of 2467 bp in length
            *   12254 12353: gap of      100 bp
            *   12354    15228: contig of 2875 bp in length
            *   15229 15328: gap of      100 bp
            *   15329    17200: contig of 1872 bp in length
            *   17201 17300: gap of      100 bp
            *   17301    20131: contig of 2831 bp in length
            *   20132 20231: gap of      100 bp
            *   20232    22587: contig of 2356 bp in length
            *   22588 22687: gap of      100 bp
            *   22688    25707: contig of 3020 bp in length
            *   25708 25807: gap of      100 bp
            *   25808    28184: contig of 2377 bp in length
            *   28185 28284: gap of      100 bp
            *   28285    31338: contig of 3054 bp in length
            *   31339 31438: gap of      100 bp
            *   31439    34299: contig of 2861 bp in length
            *   34300 34399: gap of      100 bp
            *   34400    38318: contig of 3919 bp in length
            *   38319 38418: gap of      100 bp
            *   38419    42835: contig of 4417 bp in length
            *   42836 42935: gap of      100 bp
            *   42936    45448: contig of 2513 bp in length
            *   45449 45548: gap of      100 bp
            *   45549    48116: contig of 2568 bp in length
            *   48117 48216: gap of      100 bp
            *   48217    52618: contig of 4402 bp in length
            *   52619 52718: gap of      100 bp
```

```
            *    52719    56592: contig of 3874 bp in length
            *    56593 56692: gap of      100 bp
            *    56693    59635: contig of 2943 bp in length
            *    59636 59735: gap of      100 bp
            *    59736    63661: contig of 3926 bp in length
            *    63662 63761: gap of      100 bp
            *    63762    68437: contig of 4676 bp in length
            *    68438 68537: gap of      100 bp
            *    68538    71458: contig of 2921 bp in length
            *    71459 71558: gap of      100 bp
            *    71559    76888: contig of 5330 bp in length
            *    76889 76988: gap of      100 bp
            *    76989    82113: contig of 5125 bp in length
            *    82114 82213: gap of      100 bp
            *    82214    88220: contig of 6007 bp in length
            *    88221 88320: gap of      100 bp
            *    88321    93499: contig of 5179 bp in length
            *    93500 93599: gap of      100 bp
            *    93600    97901: contig of 4302 bp in length
            *    97902 98001: gap of      100 bp
            *    98002   103016: contig of 5015 bp in length
            *   103017 103116: gap of      100 bp
            *   103117   109178: contig of 6062 bp in length
            *   109179 109278: gap of      100 bp
            *   109279   117307: contig of 8029 bp in length
            *   117308 117407: gap of      100 bp
            *   117408   124079: contig of 6672 bp in length
            *   124080 124179: gap of      100 bp
            *   124180   131281: contig of 7102 bp in length
            *   131282 131381: gap of      100 bp
            *   131382   138059: contig of 6678 bp in length
            *   138060 138159: gap of      100 bp
            *   138160   145491: contig of 7332 bp in length
            *   145492 145591: gap of      100 bp
            *   145592   157391: contig of 11800 bp in length
            *   157392 157491: gap of      100 bp
            *   157492   175064: contig of 17573 bp in length.
FEATURES             Location/Qualifiers
     source          1. .175064
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="RP11-1M10"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1003
                     /note="assembly_fragment"
     misc_feature    1104. .2634
                     /note="assembly_fragment"
     misc_feature    2735. .4415
                     /note="assembly_fragment"
     misc_feature    4516. .5785
                     /note="assembly_fragment"
     misc_feature    5886. .7879
                     /note="assembly_fragment"
     misc_feature    7980. .9686
                     /note="assembly_fragment"
     misc_feature    9787. .12253
                     /note="assembly_fragment"
     misc_feature    12354. .15228
                     /note="assembly_fragment"
     misc_feature    15329. .17200
                     /note="assembly_fragment"
     misc_feature    17301. .20131
                     /note="assembly_fragment"
     misc_feature    20232. .22587
                     /note="assembly_fragment"
     misc_feature    22688. .25707
                     /note="assembly_fragment"
     misc_feature    25808. .28184
                     /note="assembly_fragment"
     misc_feature    28285. .31338
                     /note="assembly_fragment"
     misc_feature    31439. .34299
                     /note="assembly_fragment"
     misc_feature    34400. .38318
                     /note="assembly_fragment"
     misc_feature    38419. .42835
                     /note="assembly_fragment"
     misc_feature    42936. .45448
                     /note="assembly_fragment"
     misc_feature    45549. .48116
                     /note="assembly_fragment"
     misc_feature    48217. .52618
                     /note="assembly_fragment"
     misc_feature    52719. .56592
                     /note="assembly_fragment"
     misc_feature    56693. .59635
                     /note="assembly_fragment"
     misc_feature    59736. .63661
                     /note="assembly_fragment"
     misc_feature    63762. .68437
                     /note="assembly_fragment"
     misc_feature    68538. .71458
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    71559. .76888

  Query Match             37.7%;  Score 60;  DB 2;  Length 175064;
  Best Local Similarity   100.0%;  Pred. No. 7.9e-23;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 151314 TGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAAGACAAACA 151255


RESULT  28
AX092589
LOCUS       AX092589               183999 bp    DNA     linear   PAT 21-MAR-2001
DEFINITION  Sequence 1 from Patent WO0115676.
ACCESSION   AX092589
VERSION     AX092589.1  GI:13444647
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 183999)
  AUTHORS   Hayden,M.R., Brooks-Wilson,A.R., Pimstone,S.N. and Clee,S.M.
  TITLE     Compositions and methods for modulating hdl cholesterol and
            triglyceride levels
  JOURNAL   Patent: WO 0115676-A 1 08-MAR-2001;
            University of British Columbia (CA) ; Xenon Genetics Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .183999
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT    49549 a  37944 c  41170 g  54950 t    386 others
ORIGIN


  Query Match             37.7%;  Score 60;  DB 6;  Length 183999;
  Best Local Similarity   100.0%;  Pred. No. 7.9e-23;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  53328 TGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAAGACAAACA 53387


RESULT  29
AF287262
LOCUS       AF287262               201144 bp    DNA     linear   PRI 29-APR-2001
DEFINITION  Homo sapiens ATP-binding cassette 1 sub-family A member 1 (ABCA1)
            and SNAP protein genes, complete cds.
ACCESSION   AF287262
```

```
VERSION     AF287262.1  GI:13876612
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 201144)
  AUTHORS   Qiu,Y., Cavelier,L., Chiu,S., Yang,X., Rubin,E. and Cheng,J.F.
  TITLE     Human and mouse abca1 comparative sequencing and transgenesis
            studies revealing novel regulatory sequences
  JOURNAL   Genomics 73 (1), 66-76 (2001)
  MEDLINE   21251004
REFERENCE   2  (bases 1 to 201144)
  AUTHORS   Qiu,Y., Cavelier,L., Chiu,S., Rubin,E. and Cheng,J.-F.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JUL-2000) Genome Science Department, Lawrence
            Berkeley National Laboratory, 1 Cyclotron Rd, MS 84-171, Berkeley,
            CA 94720, USA
FEATURES             Location/Qualifiers
     source          1. .201144
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     mRNA            join(33931. .34151,58319. .58476,72893. .72986,77520. .77661,
                     78931. .79049,101048. .101169,104152. .104328,117287. .117379,
                     122329. .122569,125285. .125424,125754. .125870,
                     130079. .130276,131023. .131228,131749. .131925,
                     133708. .133930,135677. .135898,136959. .137163,
                     138268. .138381,140179. .140350,141340. .141471,
                     142777. .142919,143123. .143260,143963. .144183,
                     145441. .145513,146657. .146859,148527. .148575,
                     148771. .148884,150280. .150428,152078. .152202,
                     153438. .153536,156568. .156757,158278. .158372,
                     159682. .159714,160837. .160942,162417. .162491,
                     163013. .163182,164413. .164590,166588. .166703,
                     166815. .166959,168505. .168628,169715. .169844,
                     170109. .170229,171017. .171079,171987. .172093,
                     174448. .174589,174961. .175095,176039. .176142,
                     176625. .176717,177376. .177619,178560. .181457)
                     /gene="ABCA1"
                     /product="ATP-binding cassette 1 sub-family A member 1"
     gene            33931. .181457
                     /gene="ABCA1"
     exon            33931. .34151
                     /gene="ABCA1"
                     /number=1
     exon            58319. .58476
                     /gene="ABCA1"
                     /number=2
     CDS             join(58411. .58476,72893. .72986,77520. .77661,78931. .79049,
                     101048. .101169,104152. .104328,117287. .117379,
                     122329. .122569,125285. .125424,125754. .125870,
                     130079. .130276,131023. .131228,131749. .131925,
                     133708. .133930,135677. .135898,136959. .137163,
                     138268. .138381,140179. .140350,141340. .141471,
                     142777. .142919,143123. .143260,143963. .144183,
                     145441. .145513,146657. .146859,148527. .148575,
                     148771. .148884,150280. .150428,152078. .152202,
                     153438. .153536,156568. .156757,158278. .158372,
                     159682. .159714,160837. .160942,162417. .162491,
                     163013. .163182,164413. .164590,166588. .166703,
                     166815. .166959,168505. .168628,169715. .169844,
                     170109. .170229,171017. .171079,171987. .172093,
                     174448. .174589,174961. .175095,176039. .176142,
                     176625. .176717,177376. .177619,178560. .178700)
                     /gene="ABCA1"
                     /codon_start=1
                     /product="ATP-binding cassette 1 sub-family A member 1"
                     /protein_id="AAK43526.1"
                     /db_xref="GI:13876613"
                     /translation="MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRL
                     SYPPYEQHECHFPNKAMPSAGTLPWVQGIICNANNPCFRYPTPGEAPGVVGNFNKSIV
                     ARLFSDARRLLLYSQKDTSMKDMRKVLRTLQQIKKSSSNLKLQDFLVDNETFSGFLYH
                     NLSLPKSTVDKMLRADVILHKVFLQGYQLHLTSLCNGSKSEEMIQLGDQEVSELCGLP
                     KEKLAAAERVLRSNMDILKPILRTLNSTSPFPSKELAEATKTLLHSLGTLAQELFSMR
                     SWSDMRQEVMFLTNVNSSSSSTQIYQAVSRIVCGHPEGGGLKIKSLNWYEDNNYKALF
                     GGNGTEEDAETFYDNSTTPYCNDLMKNLESSPLSRIIWKALKPLLVGKILYTPDTPAT
                     RQVMAEVNKTFQELAVFHDLEGMWEELSPKIWTFMENSQEMDLVRMLLDSRDNDHFWE
                     QQLDGLDWTAQDIVAFLAKHPEDVQSSNGSVYTWREAFNETNQAIRTISRFMECVNLN
                     KLEPIATEVWLINKSMELLDERKFWAGIVFTGITPGSIELPHHVKYKIRMDIDNVERT
                     NKIKDGYWDPGPRADPFEDMRYVWGGFAYLQDVVEQAIIRVLTGTEKKTGVYMQQMPY
                     PCYVDDIFLRVMSRSMPLFMTLAWIYSVAVIIKGIVYEKEARLKETMRIMGLDNSILW
                     FSWFISSLIPLLVSAGLLVVILKLGNLLPYSDPSVVFVFLSVFAVVTILQCFLISTLF
                     SRANLAAACGGIIYFTLYLPYVLCVAWQDYVGFTLKIFASLLSPVAFGFGCECFALFE
                     EQGIGVQWDNLFESPVEEDGFNLTTSVSMMLFNTFLYGVMTWYIEAVFPGQYGIPRPW
                     YFPCTKSYWFGEESDEKSHPGSNQKRISEICMEEEPTHLKLGVSIQNLVKVYRDGMKV
                     AVDGLALNFYEGQITSFLGHNGAGKTTTMSILTGLFPPTSGTAYILGKDIRSEMSTIR
                     QNLGVCPQHNVLFDMLTVEEHIWFYARLKGLSKKHVKAEMEQMALDVGLPSSKLKSKT
                     SQLSGGMQRKLSVALAFVGGSKVVILDEPTAGVDPYSRRGIWELLLKYRQGRTIILST
                     HHMDEADVLGDRIAIISHGKLCCVGSSLFLKNQLGTGYYLTLVKKDVESSLSSCRNSS
                     STVSYLKKEDSVSQSSSDAGLGSDHESDTLTIDVSAISNLIRKHVSEARLVEDIGHEL
                     TYVLPYEAAKEGAFVELFHEIDDRLSDLGISSYGISETTLEEIFLKVAEESGVDAETS
                     DGTLPARRNRRAFGDKQSCLRPFTEDDAADPNDSDIDPESRETDLLSGMDGKGSYQVK
                     GWKLTQQQFVALLWKRLLIARRSRKGFFAQIVLPAVFVCIALVFSLIVPPFGKYPSLE
                     LQPWMYNEQYTFVSNDAPEDTGTLELLNALTKDPGFGTRCMEGNPIPDTPCQAGEEEW
                     TTAPVPQTIMDLFQNGNWTMQNPSPACQCSSDKIKKMLPVCPPGAGGLPPPQRKQNTA
                     DILQDLTGRNISDYLVKTYVQIIAKSLKNKIWVNEFRYGGFSLGVSNTQALPPSQEVN
                     DAIKQMKKHLKLAKDSSADRFLNSLGRFMTGLDTRNNVKVWFNNKGWHAISSFLNVIN
                     NAILRANLQKGENPSHYGITAFNHPLNLTKQQLSEVALMTTSVDVLVSICVIFAMSFV
                     PASFVVFLIQERVSKAKHLQFISGVKPVIYWLSNFVWDMCNYVVPATLVIIIFICFQQ
                     KSYVSSTNLPVLALLLLLYGWSITPLMYPASFVFKIPSTAYVVLTSVNLFIGINGSVA
                     TFVLELFTDNKLNNINDILKSVFLIFPHFCLGRGLIDMVKNQAMADALERFGENRFVS
                     PLSWDLVGRNLFAMAVEGVVFFLITVLIQYRFFIRPRPVNAKLSPLNDEDEDVRRERQ
                     RILDGGGQNDILEIKELTKIYRRKRKPAVDRICVGIPPGECFGLLGVNGAGKSSTFKM
                     LTGDTTVTRGDAFLNKNSILSNIHEVHQNMGYCPQFDAITELLTGREHVEFFALLRGV
                     PEKEVGKVGEWAIRKLGLVKYGEKYAGNYSGGNKRKLSTAMALIGGPPVVFLDEPTTG
                     MDPKARRFLWNCALSVVKEGRSVVLTSHSMEECEALCTRMAIMVNGRFRCLGSVQHLK
                     NRFGDGYTIVVRIAGSNPDLKPVQDFFGLAFPGSVLKEKHRNMLQYQLPSSLSSLARI
                     FSILSQSKKRLHIEDYSVSQTTLDQVFVNFAKDQSDDDHLKDLSLHKNQTVVDVAVLT
                     SFLQDEKVKESYV"
     exon            72893. .72986
                     /gene="ABCA1"
                     /number=3
     exon            77520. .77661
                     /gene="ABCA1"
                     /number=4
     exon            78931. .79049
                     /gene="ABCA1"
                     /number=5
     exon            101048. .101169
                     /gene="ABCA1"
                     /number=6
     exon            104152. .104328
                     /gene="ABCA1"
                     /number=7
     exon            117287. .117379
                     /gene="ABCA1"
                     /number=8
     exon            122329. .122569
                     /gene="ABCA1"
                     /number=9
     exon            125285. .125424
                     /gene="ABCA1"
                     /number=10
     exon            125754. .125870
                     /gene="ABCA1"
                     /number=11
     exon            130079. .130276
                     /gene="ABCA1"
                     /number=12
     exon            131023. .131228
                     /gene="ABCA1"
                     /number=13
     exon            131749. .131925
                     /gene="ABCA1"
                     /number=14
     exon            133708. .133930
```

```
                     /gene="ABCA1"
                     /number=15
     exon            135677. .135898
                     /gene="ABCA1"
                     /number=16
     exon            136959. .137163
                     /gene="ABCA1"
                     /number=17
     exon            138268. .138381
                     /gene="ABCA1"
                     /number=18
     exon            140179. .140350
                     /gene="ABCA1"
                     /number=19
     exon            141340. .141471
                     /gene="ABCA1"
                     /number=20
     exon            142777. .142919
                     /gene="ABCA1"
                     /number=21
     exon            143123. .143260
                     /gene="ABCA1"
                     /number=22
     exon            143963. .144183
                     /gene="ABCA1"
                     /number=23
     exon            145441. .145513
                     /gene="ABCA1"
                     /number=24
     exon            146657. .146859
                     /gene="ABCA1"
                     /number=25
     exon            148527. .148575
                     /gene="ABCA1"
                     /number=26
     exon            148771. .148884
                     /gene="ABCA1"
                     /number=27
     exon            150280. .150428
                     /gene="ABCA1"
                     /number=28
     exon            152078. .152202
                     /gene="ABCA1"
                     /number=29
     exon            153438. .153536
                     /gene="ABCA1"
                     /number=30
     exon            156568. .156757
                     /gene="ABCA1"
                     /number=31

  Query Match             37.7%;  Score 60;  DB 9;  Length 201144;
  Best Local Similarity   100.0%;  Pred. No. 7.9e-23;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   58417 TGTTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAAGACAAACA 58476


RESULT  30
AK024328
LOCUS       AK024328               1556 bp    mRNA    linear   PRI 29-SEP-2000
DEFINITION  Homo sapiens cDNA FLJ14266 fis, clone PLACE1002437, highly similar
            to ATP-BINDING CASSETTE TRANSPORTER 1.
ACCESSION   AK024328
VERSION     AK024328.1  GI:10436685
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens placenta cDNA to mRNA, clone_lib:PLACE1
            clone:PLACE1002437.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
            Nishikawa,T., Nagai,K., Sugano,S., Takahashi-Fujii,A., Hara,H.,
            Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K.,
            Arita,M., Nabekura,T., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J.,
            Wakamatsu,A., Nakamura,Y., Nagahari,K., Masuho,Y. and Oshima,A.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished (2000)
REFERENCE   2  (bases 1 to 1556)
  AUTHORS   Isogai,T. and Otsuki,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,
            Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing and clone selection:
            Helix Research Institute (supported by Japan Key Technology Center
            etc.) and Department of Virology, Institute of Medical Science,
            University of Tokyo.
FEATURES             Location/Qualifiers
     source          1. .1556
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="PLACE1002437"
                     /tissue_type="placenta"
                     /clone_lib="PLACE1"
                     /note="cloning vector: pME18SFL3"
     CDS             314. .1405
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAB14887.1"
                     /db_xref="GI:10436686"
                     /translation="MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRL
                     SYPPYEQHECHFPNKAMPSAGTLPWVQGIICNANNPCFRYPTPGEAPGVVGNFNKSIV
                     ARLFSDARRLLLYSQKDTSMKDMRKVLRTLQQIKKSSSNLKLQDFLVDNETFSGFLYH
                     NLSLPKSTVDKMLRADVILHKVFLQGYQLHLTSLCNGSKSEEMIQLGDQEVSELCGLP
                     KEKLAAAERVLRSNMDILKPILMDVACDDIAHGQLTVPRSAAVAATGDAKPNMMGRET
                     LLSICASVPKVEFHERHILEHFSFCVCVSVSLFPAKGIVSFSWASFRIWVLWKAVFWQ
                     HGESMAVWEGQLGLGLNIAFEYFTSIDVG"
BASE COUNT      380 a    363 c    399 g    414 t
ORIGIN


  Query Match             32.1%;  Score 51;  DB 9;  Length 1556;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-17;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     109 cagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     329 CAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAAGACAAACA 379


RESULT  31
AC021345/c
LOCUS       AC021345               90698 bp    DNA     linear   HTG 13-JUL-2000
DEFINITION  Homo sapiens clone RP11-24J9, LOW-PASS SEQUENCE SAMPLING.
ACCESSION   AC021345
VERSION     AC021345.2  GI:9130845
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 90698)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-24J9
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 90698)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
```

Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

TITLE Direct Submission

JOURNAL Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT On Jul 13, 2000 this sequence version replaced gi:6705761.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
---------- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
---------- Project Information
Center project name: L4483
Center clone name: 24_J_9
----------.
* NOTE: This record contains 92 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 910: contig of 910 bp in length
* 911 1010: gap of 100 bp
* 1011 1873: contig of 863 bp in length
* 1874 1973: gap of 100 bp
* 1974 2824: contig of 851 bp in length
* 2825 2924: gap of 100 bp
* 2925 3802: contig of 878 bp in length
* 3803 3902: gap of 100 bp
* 3903 4816: contig of 914 bp in length
* 4817 4916: gap of 100 bp
* 4917 5759: contig of 843 bp in length
* 5760 5859: gap of 100 bp
* 5860 6764: contig of 905 bp in length
* 6765 6864: gap of 100 bp
* 6865 7747: contig of 883 bp in length
* 7748 7847: gap of 100 bp
* 7848 8755: contig of 908 bp in length
* 8756 8855: gap of 100 bp
* 8856 9753: contig of 898 bp in length
* 9754 9853: gap of 100 bp
* 9854 10757: contig of 904 bp in length
* 10758 10857: gap of 100 bp
* 10858 11732: contig of 875 bp in length
* 11733 11832: gap of 100 bp
* 11833 12739: contig of 907 bp in length
* 12740 12839: gap of 100 bp
* 12840 13710: contig of 871 bp in length
* 13711 13810: gap of 100 bp
* 13811 14684: contig of 874 bp in length
* 14685 14784: gap of 100 bp
* 14785 15662: contig of 878 bp in length
* 15663 15762: gap of 100 bp
* 15763 16677: contig of 915 bp in length
* 16678 16777: gap of 100 bp
* 16778 17678: contig of 901 bp in length
* 17679 17778: gap of 100 bp
* 17779 18679: contig of 901 bp in length
* 18680 18779: gap of 100 bp
* 18780 19632: contig of 853 bp in length
* 19633 19732: gap of 100 bp
* 19733 20634: contig of 902 bp in length
* 20635 20734: gap of 100 bp
* 20735 21620: contig of 886 bp in length
* 21621 21720: gap of 100 bp
* 21721 22579: contig of 859 bp in length
* 22580 22679: gap of 100 bp
* 22680 23568: contig of 889 bp in length
* 23569 23668: gap of 100 bp
* 23669 24554: contig of 886 bp in length
* 24555 24654: gap of 100 bp
* 24655 25521: contig of 867 bp in length
* 25522 25621: gap of 100 bp
* 25622 26487: contig of 866 bp in length
* 26488 26587: gap of 100 bp
* 26588 27464: contig of 877 bp in length
* 27465 27564: gap of 100 bp
* 27565 28466: contig of 902 bp in length
* 28467 28566: gap of 100 bp
* 28567 29464: contig of 898 bp in length
* 29465 29564: gap of 100 bp
* 29565 30447: contig of 883 bp in length
* 30448 30547: gap of 100 bp
* 30548 31453: contig of 906 bp in length
* 31454 31553: gap of 100 bp
* 31554 32452: contig of 899 bp in length
* 32453 32552: gap of 100 bp
* 32553 33447: contig of 895 bp in length
* 33448 33547: gap of 100 bp
* 33548 34435: contig of 888 bp in length
* 34436 34535: gap of 100 bp
* 34536 35433: contig of 898 bp in length
* 35434 35533: gap of 100 bp
* 35534 36440: contig of 907 bp in length
* 36441 36540: gap of 100 bp
* 36541 37422: contig of 882 bp in length
* 37423 37522: gap of 100 bp
* 37523 38402: contig of 880 bp in length
* 38403 38502: gap of 100 bp
* 38503 39380: contig of 878 bp in length
* 39381 39480: gap of 100 bp
* 39481 40407: contig of 927 bp in length
* 40408 40507: gap of 100 bp
* 40508 41405: contig of 898 bp in length
* 41406 41505: gap of 100 bp
* 41506 42399: contig of 894 bp in length
* 42400 42499: gap of 100 bp
* 42500 43357: contig of 858 bp in length
* 43358 43457: gap of 100 bp
* 43458 44356: contig of 899 bp in length
* 44357 44456: gap of 100 bp
* 44457 45325: contig of 869 bp in length
* 45326 45425: gap of 100 bp
* 45426 46305: contig of 880 bp in length
* 46306 46405: gap of 100 bp
* 46406 47302: contig of 897 bp in length
* 47303 47402: gap of 100 bp
* 47403 48293: contig of 891 bp in length
* 48294 48393: gap of 100 bp
* 48394 49280: contig of 887 bp in length
* 49281 49380: gap of 100 bp
* 49381 50250: contig of 870 bp in length
* 50251 50350: gap of 100 bp
* 50351 51209: contig of 859 bp in length
* 51210 51309: gap of 100 bp
* 51310 52221: contig of 912 bp in length

```
          *    52222 52321: gap of      100 bp
          *    52322    53205: contig of 884 bp in length
          *    53206 53305: gap of      100 bp
          *    53306    54225: contig of 920 bp in length
          *    54226 54325: gap of      100 bp
          *    54326    55206: contig of 881 bp in length
          *    55207 55306: gap of      100 bp
          *    55307    56165: contig of 859 bp in length
          *    56166 56265: gap of      100 bp
          *    56266    57124: contig of 859 bp in length
          *    57125 57224: gap of      100 bp
          *    57225    58121: contig of 897 bp in length
          *    58122 58221: gap of      100 bp
          *    58222    59075: contig of 854 bp in length
          *    59076 59175: gap of      100 bp
          *    59176    60058: contig of 883 bp in length
          *    60059 60158: gap of      100 bp
          *    60159    61067: contig of 909 bp in length
          *    61068 61167: gap of      100 bp
          *    61168    62051: contig of 884 bp in length
          *    62052 62151: gap of      100 bp
          *    62152    63022: contig of 871 bp in length
          *    63023 63122: gap of      100 bp
          *    63123    64026: contig of 904 bp in length
          *    64027 64126: gap of      100 bp
          *    64127    65015: contig of 889 bp in length
          *    65016 65115: gap of      100 bp
          *    65116    65975: contig of 860 bp in length
          *    65976 66075: gap of      100 bp
          *    66076    66954: contig of 879 bp in length
          *    66955 67054: gap of      100 bp
          *    67055    67947: contig of 893 bp in length
          *    67948 68047: gap of      100 bp
          *    68048    68934: contig of 887 bp in length
          *    68935 69034: gap of      100 bp
          *    69035    69910: contig of 876 bp in length
          *    69911 70010: gap of      100 bp

  Query Match             27.7%;  Score 44;  DB 2;  Length 90698;
  Best Local Similarity   100.0%;  Pred. No. 8.7e-14;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      32 agctctggccgctgccttccagggctcccgagccacacgctggg 75
           |||||||||||||||||||||||||||||||||||||||||||
Db   13040 AGCTCTGGCCGCTGCCTTCCAGGGCTCCCGAGCCACACGCTGGG 12997


RESULT  32
AX092843
LOCUS       AX092843                37 bp    DNA     linear   PAT 21-MAR-2001
DEFINITION  Sequence 255 from Patent WO0115676.
ACCESSION   AX092843
VERSION     AX092843.1  GI:13444900
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 37)
  AUTHORS   Hayden,M.R., Brooks-Wilson,A.R., Pimstone,S.N. and Clee,S.M.
  TITLE     Compositions and methods for modulating hdl cholesterol and
            triglyceride levels
  JOURNAL   Patent: WO 0115676-A 255 08-MAR-2001;
            University of British Columbia (CA) ; Xenon Genetics Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .37
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic Primer"
BASE COUNT        4 a      17 c      11 g       5 t
ORIGIN


  Query Match             22.0%;  Score 35;  DB 6;  Length 37;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-08;
  Matches   35;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      39 gccgctgccttccagggctcccgagccacacgctg 73
           |||||||||||||||||||||||||||||||||||
Db       1 GCCGCTGCCTTCCAGGGCTCCCGAGCCACACGCTG 35


RESULT  33
AX092705
LOCUS       AX092705                21 bp    DNA     linear   PAT 21-MAR-2001
DEFINITION  Sequence 117 from Patent WO0115676.
ACCESSION   AX092705
VERSION     AX092705.1  GI:13444762
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Hayden,M.R., Brooks-Wilson,A.R., Pimstone,S.N. and Clee,S.M.
  TITLE     Compositions and methods for modulating hdl cholesterol and
            triglyceride levels
  JOURNAL   Patent: WO 0115676-A 117 08-MAR-2001;
            University of British Columbia (CA) ; Xenon Genetics Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT        3 a      10 c       6 g       2 t
ORIGIN


  Query Match             13.2%;  Score 21;  DB 6;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 accagccacgggcgtccctgc 27
           |||||||||||||||||||||
Db       1 ACCAGCCACGGGCGTCCCTGC 21


RESULT  34
AX092707
LOCUS       AX092707                21 bp    DNA     linear   PAT 21-MAR-2001
DEFINITION  Sequence 119 from Patent WO0115676.
ACCESSION   AX092707
VERSION     AX092707.1  GI:13444764
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Hayden,M.R., Brooks-Wilson,A.R., Pimstone,S.N. and Clee,S.M.
  TITLE     Compositions and methods for modulating hdl cholesterol and
            triglyceride levels
  JOURNAL   Patent: WO 0115676-A 119 08-MAR-2001;
            University of British Columbia (CA) ; Xenon Genetics Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT        2 a       6 c       9 g       4 t
ORIGIN


  Query Match             13.2%;  Score 21;  DB 6;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      66 acacgctgggcgtgctggctg 86
           |||||||||||||||||||||
```

```
Db      1 ACACGCTGGGCGTGCTGGCTG 21


RESULT  35
AC007388/c
LOCUS       AC007388              151961 bp    DNA     linear   PRI 28-SEP-1999
DEFINITION  Homo sapiens clone NH0395B14, complete sequence.
ACCESSION   AC007388
VERSION     AC007388.3  GI:5931452
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 151961)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 151961)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   3  (bases 1 to 151961)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-SEP-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Sep 28, 1999 this sequence version replaced gi:5103896.
FEATURES             Location/Qualifiers
     source          1. .151961
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NH0395B14"
BASE COUNT    47610 a  31316 c  31453 g  41582 t
ORIGIN


  Query Match           12.6%;  Score 20;  DB 9;  Length 151961;
  Best Local Similarity  100.0%;  Pred. No. 3;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    138 cactttcagaagaagacaaa 157
          ||||||||||||||||||||
Db   9177 CACTTTCAGAAGAAGACAAA 9158


RESULT  36
AR066487
LOCUS       AR066487               624 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 9 from patent US 5850020.
ACCESSION   AR066487
VERSION     AR066487.1  GI:5996703
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 624)
  AUTHORS   Bloksberg,L.N., Havukkala,I. and Grierson,A.
  TITLE     Materials and method for the modification of plant lignin content
  JOURNAL   Patent: US 5850020-A 9 15-DEC-1998;
FEATURES             Location/Qualifiers
     source          1. .624
                     /organism="unknown"
BASE COUNT      136 a    188 c    188 g    111 t      1 others
ORIGIN


  Query Match           11.9%;  Score 19;  DB 6;  Length 624;
  Best Local Similarity  100.0%;  Pred. No. 13;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     70 gctgggcgtgctggctgag 88
          ||||||||||||||||||
Db    261 GCTGGGCGTGCTGGCTGAG 279


RESULT  37
AR074100
LOCUS       AR074100               624 bp    DNA     linear   PAT 28-AUG-2000
DEFINITION  Sequence 9 from patent US 5952486.
ACCESSION   AR074100
VERSION     AR074100.1  GI:10000860
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 624)
  AUTHORS   Bloksberg,L.N., Havukkala,I. and Grierson,A.W.
  TITLE     Materials and methods for the modification of plant lignin content
  JOURNAL   Patent: US 5952486-A 9 14-SEP-1999;
FEATURES             Location/Qualifiers
     source          1. .624
                     /organism="unknown"
BASE COUNT      136 a    188 c    188 g    111 t      1 others
ORIGIN


  Query Match           11.9%;  Score 19;  DB 6;  Length 624;
  Best Local Similarity  100.0%;  Pred. No. 13;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     70 gctgggcgtgctggctgag 88
          ||||||||||||||||||
Db    261 GCTGGGCGTGCTGGCTGAG 279


RESULT  38
AR143612
LOCUS       AR143612               624 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 9 from patent US 6204434.
ACCESSION   AR143612
VERSION     AR143612.1  GI:15104898
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 624)
  AUTHORS   Bloksberg,L.N., Havukkala,I. and Grierson,A.
  TITLE     Materials and methods for the modification of plant lignin content
  JOURNAL   Patent: US 6204434-A 9 20-MAR-2001;
FEATURES             Location/Qualifiers
     source          1. .624
                     /organism="unknown"
BASE COUNT      136 a    188 c    188 g    111 t      1 others
ORIGIN


  Query Match           11.9%;  Score 19;  DB 6;  Length 624;
  Best Local Similarity  100.0%;  Pred. No. 13;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     70 gctgggcgtgctggctgag 88
          ||||||||||||||||||
Db    261 GCTGGGCGTGCTGGCTGAG 279


RESULT  39
BD005648
LOCUS       BD005648               624 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Materials and methods for the modification of plant lignin content.
ACCESSION   BD005648
VERSION     BD005648.1  GI:18634019
KEYWORDS    JP 2001500378-A/9.
```

```
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 624)
  AUTHORS   Bloksberg,L.N., Grierson,A.W. and Havukkala,I.J.
  TITLE     Materials and methods for the modification of plant lignin content
  JOURNAL   Patent: JP 2001500378-A 9 16-JAN-2001;
            GENESIS RESEARCH & DEVELOPMENT CO LTD, LETCHER CHALLENGE FORESTS
            LTD
COMMENT     OS   Unidentified
            PN   JP 2001500378-A/9
            PD   16-JAN-2001
            PF   10-SEP-1997 JP 1998513535
            PR   11-SEP-1996 US    08/713000
            PI   LEONARD NATHAN BLOKSBERG,ALISTAIR WALLACE GRIERSON, PI   ILKKA
            JAAKKO HAVUKKALA
            PC   C12N15/53,C12N15/54,C12N15/52,C12N15/60,C12N15/82,A01H5/00 CC
             Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .624
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .624
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT      136 a    188 c    188 g    111 t      1 others
ORIGIN


  Query Match             11.9%;  Score 19;  DB 6;  Length 624;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      70 gctgggcgtgctggctgag 88
           ||||||||||||||||||||
Db     261 GCTGGGCGTGCTGGCTGAG 279


RESULT  40
AR074136
LOCUS       AR074136                684 bp    DNA     linear   PAT 28-AUG-2000
DEFINITION  Sequence 45 from patent US 5952486.
ACCESSION   AR074136
VERSION     AR074136.1  GI:10000896
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 684)
  AUTHORS   Bloksberg,L.N., Havukkala,I. and Grierson,A.W.
  TITLE     Materials and methods for the modification of plant lignin content
  JOURNAL   Patent: US 5952486-A 45 14-SEP-1999;
FEATURES             Location/Qualifiers
     source          1. .684
                     /organism="unknown"
BASE COUNT      150 a    207 c    200 g    127 t
ORIGIN


  Query Match             11.9%;  Score 19;  DB 6;  Length 684;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      70 gctgggcgtgctggctgag 88
           ||||||||||||||||||||
Db     261 GCTGGGCGTGCTGGCTGAG 279


RESULT  41
BD005684
LOCUS       BD005684                684 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Materials and methods for the modification of plant lignin content.
ACCESSION   BD005684
VERSION     BD005684.1  GI:18634055
KEYWORDS    JP 2001500378-A/45.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 684)
  AUTHORS   Bloksberg,L.N., Grierson,A.W. and Havukkala,I.J.
  TITLE     Materials and methods for the modification of plant lignin content
  JOURNAL   Patent: JP 2001500378-A 45 16-JAN-2001;
            GENESIS RESEARCH & DEVELOPMENT CO LTD, LETCHER CHALLENGE FORESTS
            LTD
COMMENT     OS   Unidentified
            PN   JP 2001500378-A/45
            PD   16-JAN-2001
            PF   10-SEP-1997 JP 1998513535
            PR   11-SEP-1996 US    08/713000
            PI   LEONARD NATHAN BLOKSBERG,ALISTAIR WALLACE GRIERSON, PI   ILKKA
            JAAKKO HAVUKKALA
            PC   C12N15/53,C12N15/54,C12N15/52,C12N15/60,C12N15/82,A01H5/00 CC
             Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .684
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .684
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT      150 a    207 c    200 g    127 t
ORIGIN


  Query Match             11.9%;  Score 19;  DB 6;  Length 684;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      70 gctgggcgtgctggctgag 88
           ||||||||||||||||||||
Db     261 GCTGGGCGTGCTGGCTGAG 279


RESULT  42
AB014524
LOCUS       AB014524                6542 bp    mRNA    linear   PRI 06-FEB-1999
DEFINITION  Homo sapiens mRNA for KIAA0624 protein, partial cds.
ACCESSION   AB014524
VERSION     AB014524.1  GI:3327061
KEYWORDS    .
SOURCE      Homo sapiens adult male brain cDNA to mRNA, clone_lib:pBluescriptII
            SK plus clone:HG04767.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 6542)
  AUTHORS   Ohara,O., Suyama,M., Nagase,T. and Ishikawa,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-MAY-1998) Osamu Ohara, Kazusa DNA Research Institute,
            Laboratory of DNA Technology; Yana 1532-3, Kisarazu, Chiba
            292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, Tel:+81-438-52-3913,
            Fax:+81-438-52-3914)
REFERENCE   2  (sites)
  AUTHORS   Ishikawa,K., Nagase,T., Suyama,M., Miyajima,N., Tanaka,A.,
            Kotani,H., Nomura,N. and Ohara,O.
  TITLE     Prediction of the coding sequences of unidentified human genes. X.
            The complete sequences of 100 new cDNA clones from brain which can
            code for large proteins in vitro
  JOURNAL   DNA Res. 5 (3), 169-176 (1998)
  MEDLINE   98403880
FEATURES             Location/Qualifiers
     source          1. .6542
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
```

```
                     /clone="HG04767"
                     /sex="male"
                     /tissue_type="brain"
                     /clone_lib="pBluescriptII SK plus"
                     /dev_stage="adult"
     gene            1. .5954
                     /gene="KIAA0624"
     CDS             <1. .5954
                     /gene="KIAA0624"
                     /codon_start=3
                     /product="KIAA0624 protein"
                     /protein_id="BAA31599.1"
                     /db_xref="GI:3327062"
                     /translation="NMSLSFLNDEEARKILQVLERNEELQRAEKDRISKLQKTKRDIR
                     WLQGVTGEWFEEIQRKKFCNETDVSQMLKQPLTYRLSKEMAKNDPIELPTSRSKNVTN
                     QKKPTPFSSRMSFRSSFASLFSFRKSGKETSKLPSLGQKGCDGHAGPPMPVRGAAVQA
                     KIYNSPLENHLVDSTFVPKPAVMREESGMPPPWDASLLENEFFQVLDDLDSKLAQEQS
                     ASSVNTRTPLNYGSRTQFGHFYSSGNRHGNITERHKKHYNETSNMSIYDILRPGTPRE
                     GFKTFSPRTSTIYDMYRTREPRVFKEDYVQKNTFGSTSLCFDSRQRSALPATGHFTAR
                     SLHFPATTQSKSGFIPPRHQQSPKRTPLSSIIWNRSDSSRDRENQEEFLRAPSPMEID
                     PADKYVYPRGFQENKRYESYHSQNVYQRVSLNAPMENAMSPDTFENSENMPFYHQSNT
                     FTRSFFSNTFGRSGEQRRFGQGPFWGQEKGHSFWSDFHRSRKSFSSSDRDFEMISMEA
                     NSVSAIHGHNVSSEHWESFSSGYGTDVSRGQEEPHPWQFDFQRSTLDSMVVSHGNETQ
                     LTPHFGTPNVCSMTGSSYHVKSSELVSQQDSSPVEVHINKEASSFGIAQTLASSFKTS
                     FSQISDDRRNPQSPNLQNPTVTLQKIFPNKPASHPMRSHTEVTVTNSNSVDSLPLAKS
                     QPNILVTEVNNEKDLNESISEEDKQLSKMDQTNKAGEIPQPVSQTGISNSLPDFQNPL
                     SQDSAKSNGFGFNASTIISSKKSPRVFSRKDTSKMYIPHTDKSNDIKQDKRFTENRKL
                     GSTASLPFIQEHRTPPSFPRTDQGCHQELTVNNEDISRIITNNHWSSALTDTQNAQYS
                     KCKLTPGHKTSCDSLDLSSAALPDSSPSKNSSLDAPVVPSTTVFSRRSPSDKDPSLGE
                     REEKDNAGKNQKNQFIVSHSENQERNDSPVPTHDEVVDVKCHSHSPFRNERGKGKIRH
                     HISCIEKLSKTESISVPTSDHRSLIEANQSNSKVSELDTIYCTLPRKSSSFLIHGRQS
                     GSKIMAASLRNGPPPFQIKNNVEDAMGNYMLNKFSPSSPESANECSKVLSDSALEAPE
                     ATERMTNVKSSGSTSVRKGPLPFLINRAMSCPSGEPHASTGREGRKKPLTSGMDASEL
                     TPRAWERIISPVESDSSVRDCSLTKRQHQKENFQEYTEKEGKMAASRRSVFALSNEDP
                     LPFCSDLSGKERGKTLHKVKTTSTFSVSGDEDNVKCLEVVSIYYTLPRKPSKKFCNLL
                     QQYTQNTNLLIESPQVETETFPNALEKDKQNYSTREQSGTPSCENLKMSVNSDQTLTT
                     ENMTAFRLSNRGPLAPTLQEMASVEAAVSLPEEESKAREIFSDNLAKTPLGDSENKKE
                     RGKKLQSETLHTSLMLQRKNVSEEKSENCQQSINSSNSGPSSLPALSEVNIGNSQTRR
                     SSWECTGSGRAIPFTGSGKCPQKDHTSTAVGDGSSGSQPREGRGDIGTNCQKMTNKTL
                     SHSESQVFALTPALHKLQLGEETQSDEPNLESLQSEPRELPQRSQEANMTESRKAEDE
                     MQKSAWDQPSLPEGNKNKTNLDDLVKGENRSSVKHRLAAMSKASRKFPAKDVSPRRHV
                     ATIFPQSGSRSGFDHLSLGTVECNPLFPEPTPKSAESIGESRLSENRKHVKKSENLLP
                     ITVLPNREPSTHVSNQKSNSISQRHQNEFKNVSESPSKHENSKDVTAAQNLVRESGAP
                     SPITFTSLREAEFSDNQRRLSPPFPLEPAQKSRVSSPLASFLQQQRSASSLEWEPEPH
                     LYRSKSLKSINVHGDLLRKSHPPKVRERHFSESTSIDNALSRLTLGNEFSVNNGYSRR
                     FRSFSELPSCDGNESWAYRSGTKTGPRSAISIYRPIDYGIFGKEQQLAFLENVKRSLT
                     QGRLWKPSFLKNPGFLKDDLRNPPNPSESLSSNSPSSQVPEDGLSPSEPLNIYEDDPV
                     DSDCDTDTTTDDEYYLDENDKESEL"
BASE COUNT     2174 a   1428 c   1312 g   1628 t
ORIGIN


  Query Match             11.9%;  Score 19;  DB 9;  Length 6542;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     141 tttcagaagaagacaaaca 159
           ||||||||||||||||||||
Db    2104 TTTCAGAAGAAGACAAACA 2122


RESULT  43
AC022089/c
LOCUS       AC022089               44504 bp    DNA     linear   PRI 03-OCT-2001
DEFINITION  Homo sapiens chromosome 5 clone CTB-125N5, complete sequence.
ACCESSION   AC022089
VERSION     AC022089.7  GI:15887304
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 44504)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 44504)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JAN-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 44504)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-AUG-2000) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 44504)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-OCT-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Oct 3, 2001 this sequence version replaced gi:9945146.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.7% of Sequence;
            Estimated Total Number of Errors is 0.2.
FEATURES             Location/Qualifiers
     source          1. .44504
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTB-125N5"
BASE COUNT    14182 a   7969 c   8161 g  14192 t
ORIGIN


  Query Match             11.9%;  Score 19;  DB 9;  Length 44504;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      98 catgttggcctcagctgag 116
           ||||||||||||||||||||
Db  30425 CATGTTGGCCTCAGCTGAG 30407


RESULT  44
AC013605/c
LOCUS       AC013605               64041 bp    DNA     linear   HTG 13-JUL-2000
DEFINITION  Homo sapiens clone RP11-10E16, LOW-PASS SEQUENCE SAMPLING.
ACCESSION   AC013605
VERSION     AC013605.2  GI:9123816
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 64041)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-10E16
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 64041)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
```

```
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jul 13, 2000 this sequence version replaced gi:6403649.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L3053
                Center clone name: 10_E_16
            --------------.
            * NOTE: This record contains 72 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *        1      820: contig of 820 bp in length
            *      821 920: gap of      100 bp
            *      921     1730: contig of 810 bp in length
            *     1731 1830: gap of      100 bp
            *     1831     2644: contig of 814 bp in length
            *     2645 2744: gap of      100 bp
            *     2745     3547: contig of 803 bp in length
            *     3548 3647: gap of      100 bp
            *     3648     4437: contig of 790 bp in length
            *     4438 4537: gap of      100 bp
            *     4538     5340: contig of 803 bp in length
            *     5341 5440: gap of      100 bp
            *     5441     6217: contig of 777 bp in length
            *     6218 6317: gap of      100 bp
            *     6318     7106: contig of 789 bp in length
            *     7107 7206: gap of      100 bp
            *     7207     8003: contig of 797 bp in length
            *     8004 8103: gap of      100 bp
            *     8104     8893: contig of 790 bp in length
            *     8894 8993: gap of      100 bp
            *     8994     9706: contig of 713 bp in length
            *     9707 9806: gap of      100 bp
            *     9807    10597: contig of 791 bp in length
            *    10598 10697: gap of      100 bp
            *    10698    11497: contig of 800 bp in length
            *    11498 11597: gap of      100 bp
            *    11598    12413: contig of 816 bp in length
            *    12414 12513: gap of      100 bp
            *    12514    13325: contig of 812 bp in length
            *    13326 13425: gap of      100 bp
            *    13426    14213: contig of 788 bp in length
            *    14214 14313: gap of      100 bp
            *    14314    15081: contig of 768 bp in length
            *    15082 15181: gap of      100 bp
            *    15182    16001: contig of 820 bp in length
            *    16002 16101: gap of      100 bp
            *    16102    16916: contig of 815 bp in length
            *    16917 17016: gap of      100 bp
            *    17017    17840: contig of 824 bp in length
            *    17841 17940: gap of      100 bp
            *    17941    18743: contig of 803 bp in length
            *    18744 18843: gap of      100 bp
            *    18844    19547: contig of 704 bp in length
            *    19548 19647: gap of      100 bp
            *    19648    20462: contig of 815 bp in length
            *    20463 20562: gap of      100 bp
            *    20563    21327: contig of 765 bp in length
            *    21328 21427: gap of      100 bp
            *    21428    22219: contig of 792 bp in length
            *    22220 22319: gap of      100 bp
            *    22320    23125: contig of 806 bp in length
            *    23126 23225: gap of      100 bp
            *    23226    24025: contig of 800 bp in length
            *    24026 24125: gap of      100 bp
            *    24126    24935: contig of 810 bp in length
            *    24936 25035: gap of      100 bp
            *    25036    25813: contig of 778 bp in length
            *    25814 25913: gap of      100 bp
            *    25914    26702: contig of 789 bp in length
            *    26703 26802: gap of      100 bp
            *    26803    27601: contig of 799 bp in length
            *    27602 27701: gap of      100 bp
            *    27702    28480: contig of 779 bp in length
            *    28481 28580: gap of      100 bp
            *    28581    29373: contig of 793 bp in length
            *    29374 29473: gap of      100 bp
            *    29474    30254: contig of 781 bp in length
            *    30255 30354: gap of      100 bp
            *    30355    31113: contig of 759 bp in length
            *    31114 31213: gap of      100 bp
            *    31214    32012: contig of 799 bp in length
            *    32013 32112: gap of      100 bp
            *    32113    32922: contig of 810 bp in length
            *    32923 33022: gap of      100 bp
            *    33023    33823: contig of 801 bp in length
            *    33824 33923: gap of      100 bp
            *    33924    34720: contig of 797 bp in length
            *    34721 34820: gap of      100 bp
            *    34821    35625: contig of 805 bp in length
            *    35626 35725: gap of      100 bp
            *    35726    36531: contig of 806 bp in length
            *    36532 36631: gap of      100 bp
            *    36632    37409: contig of 778 bp in length
            *    37410 37509: gap of      100 bp
            *    37510    38301: contig of 792 bp in length
            *    38302 38401: gap of      100 bp
            *    38402    39181: contig of 780 bp in length
            *    39182 39281: gap of      100 bp
            *    39282    40076: contig of 795 bp in length
            *    40077 40176: gap of      100 bp
            *    40177    40944: contig of 768 bp in length
            *    40945 41044: gap of      100 bp
            *    41045    41854: contig of 810 bp in length
            *    41855 41954: gap of      100 bp
            *    41955    42765: contig of 811 bp in length
            *    42766 42865: gap of      100 bp
            *    42866    43671: contig of 806 bp in length
            *    43672 43771: gap of      100 bp
            *    43772    44560: contig of 789 bp in length
            *    44561 44660: gap of      100 bp
            *    44661    45461: contig of 801 bp in length
            *    45462 45561: gap of      100 bp
            *    45562    46369: contig of 808 bp in length
            *    46370 46469: gap of      100 bp
            *    46470    47280: contig of 811 bp in length
            *    47281 47380: gap of      100 bp
            *    47381    48189: contig of 809 bp in length
            *    48190 48289: gap of      100 bp
            *    48290    49093: contig of 804 bp in length
            *    49094 49193: gap of      100 bp
            *    49194    49860: contig of 667 bp in length
            *    49861 49960: gap of      100 bp
            *    49961    50742: contig of 782 bp in length
            *    50743 50842: gap of      100 bp
            *    50843    51611: contig of 769 bp in length
            *    51612 51711: gap of      100 bp
            *    51712    52514: contig of 803 bp in length
            *    52515 52614: gap of      100 bp
```

```
            *    52615    53425: contig of 811 bp in length
            *    53426 53525: gap of      100 bp
            *    53526    54329: contig of 804 bp in length
            *    54330 54429: gap of      100 bp
            *    54430    55247: contig of 818 bp in length
            *    55248 55347: gap of      100 bp
            *    55348    56149: contig of 802 bp in length
            *    56150 56249: gap of      100 bp
            *    56250    56953: contig of 704 bp in length
            *    56954 57053: gap of      100 bp
            *    57054    57865: contig of 812 bp in length
            *    57866 57965: gap of      100 bp
            *    57966    58759: contig of 794 bp in length
            *    58760 58859: gap of      100 bp
            *    58860    59636: contig of 777 bp in length
            *    59637 59736: gap of      100 bp
            *    59737    60441: contig of 705 bp in length
            *    60442 60541: gap of      100 bp
            *    60542    61327: contig of 786 bp in length
            *    61328 61427: gap of      100 bp
            *    61428    62217: contig of 790 bp in length
            *    62218 62317: gap of      100 bp
            *    62318    63129: contig of 812 bp in length
            *    63130 63229: gap of      100 bp
            *    63230    64041: contig of 812 bp in length.
FEATURES             Location/Qualifiers
     source          1. .64041

  Query Match             11.9%;  Score 19;  DB 2;  Length 64041;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     22 ccctgctgtcagctctggc 40
          ||||||||||||||||||
Db  22122 CCCTGCTGTCAGCTCTGGC 22104


RESULT  45
AL591133
LOCUS       AL591133               66477 bp    DNA     linear   PRI 04-OCT-2001
DEFINITION  Human DNA sequence from clone RP11-12I16 on chromosome 9, complete
            sequence.
ACCESSION   AL591133
VERSION     AL591133.7  GI:15982094
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 66477)
  AUTHORS   Sehra,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Oct 5, 2001 this sequence version replaced gi:14787626.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr9
            RP11-12I16 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            IMPORTANT: This sequence is not the entire insert of clone
            RP11-12I16 It may be shorter because we sequence overlapping
            sections only once, except for a short overlap.
            The true left end of clone RP11-141L16 is at 64478 in this
            sequence. The true right end of clone RP11-338L20 is at 2000 in
            this sequence.
FEATURES             Location/Qualifiers
     source          1. .66477
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="9"
                     /clone="RP11-12I16"
                     /clone_lib="RPCI-11.1"
BASE COUNT    21366 a  12101 c  11676 g  21334 t
ORIGIN


  Query Match             11.9%;  Score 19;  DB 9;  Length 66477;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    141 tttcagaagaagacaaaca 159
          ||||||||||||||||||
Db  19651 TTTCAGAAGAAGACAAACA 19669


Search completed: September 20, 2002, 06:33:24
Job time: 18598 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 06:08:05 ; Search time 521.76 Seconds
(without alignments)
523.209 Million cell updates/sec

Title: US-09-846-456-5
Perfect score: 159
Sequence: 1 ttaatgaccagccacgggcg.....ctttcagaagaagacaaaca 159

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	77	48.4	10442	22	AAF24680	Nucleotide sequenc
2	77	48.4	10442	22	AAF24702	Nucleotide sequenc
3	77	48.4	10474	22	AAF24685	Nucleotide sequenc
4	77	48.4	10474	22	AAF24686	Nucleotide sequenc
5	77	48.4	10474	22	AAF24707	Nucleotide sequenc
6	77	48.4	10474	22	AAF24708	Nucleotide sequenc
7	60	37.7	446	22	AAS04035	Partial human ABC1
8	60	37.7	7086	22	ABA09200	Human ABCA1 homolo
9	60	37.7	7086	22	AAK52667	Human polynucleoti
10	60	37.7	7260	22	AAD21326	Human ATP binding
11	60	37.7	7260	22	AAI70315	Human ATP binding
12	60	37.7	7281	22	AAK51683	Human polynucleoti
13	60	37.7	7857	21	AAC69388	Human ABC1 cholest
14	60	37.7	7860	22	AAF83826	Human ABC1 nucleot
15	60	37.7	7860	22	AAF92835	Human ABC1 cDNA.
16	60	37.7	7861	21	AAC69387	Human ABC1 cholest
17	60	37.7	7864	21	AAC69120	Human ABC1 cholest
18	60	37.7	7864	21	AAC69385	Human ABC1 cholest
19	60	37.7	7864	21	AAC69386	Human ABC1 cholest
20	60	37.7	7864	21	AAC69389	Human ABC1 cholest
21	60	37.7	9741	22	AAS06120	Human ABC1 DNA seq
22	60	37.7	9854	22	AAS06121	Human ABC1 DNA seq
23	60	37.7	10545	21	AAC69132	Human ABC1 gene ex
24	60	37.7	183999	22	AAF92831	Human ABC1 genomic
25	51	32.1	736	22	AAH07432	Human cDNA clone (
26	51	32.1	1556	22	AAH18606	Human cDNA sequenc
27	35	22.0	37	22	AAF93084	ABC1 polymorphism
28	21	13.2	21	21	AAC69306	Human ABC1 gene pr
29	21	13.2	21	21	AAC69308	Human ABC1 gene pr
30	21	13.2	21	22	AAF92946	Polymorphic sequen
31	21	13.2	21	22	AAF92948	Polymorphic sequen
32	19	11.9	577	21	AAA68004	Pinus radiata PAL
33	19	11.9	624	19	AAV23916	Plant PAL enzyme D
34	19	11.9	624	20	AAZ06895	Pine phenylalanine
35	19	11.9	624	21	AAA67916	Pinus radiata PAL
36	19	11.9	684	19	AAV23865	Plant PAL enzyme D
37	19	11.9	684	20	AAZ06898	Pine phenylalanine
38	19	11.9	684	21	AAA69586	Pinus radiata phen
39	19	11.9	684	21	AAA67952	Pinus radiata PAL
40	19	11.9	5286	23	AAS73156	DNA encoding novel
41	19	11.9	5954	23	AAS80591	DNA encoding novel
c 42	19	11.9	6143	23	AAS83843	DNA encoding novel
c 43	18	11.3	18	21	AAC69153	Human ABC1 phospho
c 44	18	11.3	6420	23	ABL08833	Drosophila melanog
c 45	18	11.3	11580	23	ABL08832	Drosophila melanog

ALIGNMENTS

RESULT 1
AAF24680
ID AAF24680 standard; DNA; 10442 BP.
XX
AC AAF24680;
XX
DT 20-APR-2001 (first entry)
XX
DE Nucleotide sequence of a human ABC1 polypeptide.
XX
KW Human; adenosine triphosphate binding cassette protein 1; ABC1;
KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;
KW atherosclerosis; cholesterol transport; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 291..7076
FT /*tag= a
FT /product= "ABC1 polypeptide"
XX
PN WO200078972-A2.
XX
PD 28-DEC-2000.
XX
PF 16-JUN-2000; 2000WO-US16765.
XX
PR 18-JUN-1999; 99US-0140264.
PR 14-SEP-1999; 99US-0153872.
PR 19-NOV-1999; 99US-0166573.
XX

PA (CVTH-) CV THERAPEUTICS INC.
XX
PI Lawn RM, Wade D, Garvin M;
XX
DR WPI; 2001-137812/14.
XX
PT Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,
PT useful for the development of agents for the treatment of heart disease
PT and other disorders associated with hypercholesterolemia and
PT atherosclerosis -
XX
PS Disclosure; Page 122-128; 215pp; English.
XX
CC The present sequence encodes a human adenosine triphosphate (ATP)
CC binding cassette protein (ABC) 1 polypeptide. ABC1 resides in cell
CC membranes and utilises ATP hydrolysis to transport a wide variety of
CC substrates across the plasma membrane. ABC1 is a pivotal protein in
CC the apolipoprotein-mediated mobilisation of intracellular cholesterol
CC stores. ABC1 is defective in Tangier disease, a genetic disorder
CC characterised by abnormal HDL-cholesterol metabolism. The ABC1 gene is
CC localised to chromosome 9q22-9q31. The ABC1 genes and proteins are
CC useful for developing pharmaceutical agents for the treatment of heart
CC disease and other disorders associated with hypercholesterolemia and
CC atherosclerosis. The genes are useful for developing screening assays to
CC screen for compounds that regulate the expression of genes associated
CC with cholesterol transport. The genes and proteins are also useful for
CC are also useful as diagnostic indicators of cardiovascular disease and
CC other disorders associated with hypercholesterolemia.
XX
SQ Sequence 10442 BP; 2898 A; 2297 C; 2408 G; 2835 T; 4 other;

Query Match 48.4%; Score 77; DB 22; Length 10442;
Best Local Similarity 99.2%; Pred. No. 5e-29;
Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      32 agctctggccgctgccttccagggctcccgagccacacgctgggcgtgctggctgaggga 91
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     229 agctctggccgctgccttccagggctcccgagccacacgctgggcgtgctggctgaggga 288

Qy      92 acatggcatgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaa 151
           ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     289 acatggcttgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaa 348

Qy     152 gacaaaca 159
           ||||||||
Db     349 gacaaaca 356
```

RESULT 2
AAF24702
ID AAF24702 standard; DNA; 10442 BP.
XX
AC AAF24702;
XX
DT 20-APR-2001 (first entry)
XX
DE Nucleotide sequence of a human ABC1 polypeptide.
XX
KW Human; adenosine triphosphate binding cassette protein 1; ABC1;
KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;
KW atherosclerosis; cholesterol transport; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 291..7076
FT /*tag= a
FT /product= "ABC1 polypeptide"
XX
PN WO200078971-A2.
XX
PD 28-DEC-2000.
XX
PF 16-JUN-2000; 2000WO-US16591.
XX
PR 18-JUN-1999; 99US-0140264.
PR 14-SEP-1999; 99US-0153872.
PR 19-NOV-1999; 99US-0166573.
XX
PA (CVTH-) CV THERAPEUTICS INC.
PA (UNIW) UNIV WASHINGTON.
XX
PI Lawn RM, Wade D, Oram JF, Garvin M;
XX
DR WPI; 2001-137811/14.
DR P-PSDB; AAB31365.
XX
PT Adenosine triphosphate (ATP) binding cassette protein (ABC) 1
PT polynucleotides and polypeptides, useful for treatment of heart disease
PT and other disorders associated with hypercholesterolemia and
PT atherosclerosis -
XX
PS Claim 3; Page 117-123; 211pp; English.
XX
CC The present sequence encodes a human adenosine triphosphate (ATP)
CC binding cassette protein (ABC) 1 polypeptide. ABC1 resides in cell
CC membranes and utilises ATP hydrolysis to transport a wide variety of
CC substrates across the plasma membrane. ABC1 is a pivotal protein in
CC the apolipoprotein-mediated mobilisation of intracellular cholesterol
CC stores. ABC1 is defective in Tangier disease, a genetic disorder
CC characterised by abnormal HDL-cholesterol metabolism. The ABC1 gene is
CC localised to chromosome 9q22-9q31. The ABC1 genes and proteins are
CC useful for developing pharmaceutical agents for the treatment of heart
CC disease and other disorders associated with hypercholesterolemia and
CC atherosclerosis. The genes are useful for developing screening assays to
CC screen for compounds that regulate the expression of genes associated
CC with cholesterol transport. The genes and proteins are also useful for
CC are also useful as diagnostic indicators of cardiovascular disease and
CC other disorders associated with hypercholesterolemia.
XX
SQ Sequence 10442 BP; 2898 A; 2297 C; 2408 G; 2835 T; 4 other;

Query Match 48.4%; Score 77; DB 22; Length 10442;
Best Local Similarity 99.2%; Pred. No. 5e-29;
Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      32 agctctggccgctgccttccagggctcccgagccacacgctgggcgtgctggctgaggga 91
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     229 agctctggccgctgccttccagggctcccgagccacacgctgggcgtgctggctgaggga 288

Qy      92 acatggcatgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaa 151
           ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     289 acatggcttgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaa 348

Qy     152 gacaaaca 159
           ||||||||
Db     349 gacaaaca 356
```

RESULT 3
AAF24685
ID AAF24685 standard; DNA; 10474 BP.
XX
AC AAF24685;
XX
DT 20-APR-2001 (first entry)
XX
DE Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.
XX
KW Human; adenosine triphosphate binding cassette protein 1; ABC1;
KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;
KW atherosclerosis; cholesterol transport; ss.

```
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             323..7108
FT                   /*tag=  a
FT                   /product= "defective ABC1 polypeptide"
XX
PN   WO200078972-A2.
XX
PD   28-DEC-2000.
XX
PF   16-JUN-2000; 2000WO-US16765.
XX
PR   18-JUN-1999;   99US-0140264.
PR   14-SEP-1999;   99US-0153872.
PR   19-NOV-1999;   99US-0166573.
XX
PA   (CVTH-) CV THERAPEUTICS INC.
XX
PI   Lawn RM,  Wade D,  Garvin M;
XX
DR   WPI; 2001-137812/14.
XX
PT   Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,
PT   useful for the development of agents for the treatment of heart disease
PT   and other disorders associated with hypercholesterolemia and
PT   atherosclerosis -
XX
PS   Disclosure; Page 148-154; 215pp; English.
XX
CC   The present sequence encodes a human adenosine triphosphate (ATP)
CC   binding cassette protein (ABC) 1 polypeptide, and is isolated from
CC   a Tangier disease patient. ABC1 resides in cell membranes and utilises
CC   ATP hydrolysis to transport a wide variety of substrates across the
CC   plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated
CC   mobilisation of intracellular cholesterol stores. ABC1 is defective in
CC   Tangier disease, a genetic disorder characterised by abnormal
CC   HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome
CC   9q22-9q31. The ABC1 genes and proteins are useful for developing
CC   pharmaceutical agents for the treatment of heart disease and other
CC   disorders associated with hypercholesterolemia and atherosclerosis. The
CC   genes are useful for developing screening assays to screen for compounds
CC   that regulate the expression of genes associated with cholesterol
CC   transport. The genes and proteins are also useful for are also useful
CC   as diagnostic indicators of cardiovascular disease and other disorders
CC   associated with hypercholesterolemia.
XX
SQ   Sequence 10474 BP; 2906 A; 2305 C; 2416 G; 2843 T; 4 other;


  Query Match             48.4%;  Score 77;  DB 22;  Length 10474;
  Best Local Similarity   99.2%;  Pred. No. 5e-29;
  Matches  127;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      32 agctctggccgctgccttccagggctcccgagccacacgctgggcgtgctggctgaggga 91
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     261 agctctggccgctgccttccagggctcccgagccacacgctgggcgtgctggctgaggga 320

Qy      92 acatggcatgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaa 151
           ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     321 acatggcttgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaa 380

Qy     152 gacaaaca 159
           ||||||||
Db     381 gacaaaca 388


RESULT   4
AAF24686
ID   AAF24686 standard; DNA; 10474 BP.
XX
AC   AAF24686;
XX
DT   20-APR-2001  (first entry)
XX
DE   Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.
XX
KW   Human; adenosine triphosphate binding cassette protein 1; ABC1;
KW   apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW   chromosome 9q22-9q31; heart disease; hypercholesterolemia;
KW   atherosclerosis; cholesterol transport; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             323..7108
FT                   /*tag=  a
FT                   /product= "defective ABC1 polypeptide"
XX
PN   WO200078972-A2.
XX
PD   28-DEC-2000.
XX
PF   16-JUN-2000; 2000WO-US16765.
XX
PR   18-JUN-1999;   99US-0140264.
PR   14-SEP-1999;   99US-0153872.
PR   19-NOV-1999;   99US-0166573.
XX
PA   (CVTH-) CV THERAPEUTICS INC.
XX
PI   Lawn RM,  Wade D,  Garvin M;
XX
DR   WPI; 2001-137812/14.
XX
PT   Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,
PT   useful for the development of agents for the treatment of heart disease
PT   and other disorders associated with hypercholesterolemia and
PT   atherosclerosis -
XX
PS   Disclosure; Page 170-176; 215pp; English.
XX
CC   The present sequence encodes a human adenosine triphosphate (ATP)
CC   binding cassette protein (ABC) 1 polypeptide, and is isolated from
CC   a Tangier disease patient. ABC1 resides in cell membranes and utilises
CC   ATP hydrolysis to transport a wide variety of substrates across the
CC   plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated
CC   mobilisation of intracellular cholesterol stores. ABC1 is defective in
CC   Tangier disease, a genetic disorder characterised by abnormal
CC   HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome
CC   9q22-9q31. The ABC1 genes and proteins are useful for developing
CC   pharmaceutical agents for the treatment of heart disease and other
CC   disorders associated with hypercholesterolemia and atherosclerosis. The
CC   genes are useful for developing screening assays to screen for compounds
CC   that regulate the expression of genes associated with cholesterol
CC   transport. The genes and proteins are also useful for are also useful
CC   as diagnostic indicators of cardiovascular disease and other disorders
CC   associated with hypercholesterolemia.
XX
SQ   Sequence 10474 BP; 2907 A; 2304 C; 2415 G; 2844 T; 4 other;


  Query Match             48.4%;  Score 77;  DB 22;  Length 10474;
  Best Local Similarity   99.2%;  Pred. No. 5e-29;
  Matches  127;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      32 agctctggccgctgccttccagggctcccgagccacacgctgggcgtgctggctgaggga 91
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     261 agctctggccgctgccttccagggctcccgagccacacgctgggcgtgctggctgaggga 320

Qy      92 acatggcatgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaa 151
           ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     321 acatggcttgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaa 380

Qy     152 gacaaaca 159
```

```
          ||||||||
Db    381 gacaaaca 388


RESULT   5
AAF24707
ID   AAF24707 standard; DNA; 10474 BP.
XX
AC   AAF24707;
XX
DT   20-APR-2001  (first entry)
XX
DE   Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.
XX
KW   Human; adenosine triphosphate binding cassette protein 1; ABC1;
KW   apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW   chromosome 9q22-9q31; heart disease; hypercholesterolemia;
KW   atherosclerosis; cholesterol transport; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             323..7108
FT                   /*tag=  a
FT                   /product= "defective ABC1 polypeptide"
XX
PN   WO200078971-A2.
XX
PD   28-DEC-2000.
XX
PF   16-JUN-2000; 2000WO-US16591.
XX
PR   18-JUN-1999;   99US-0140264.
PR   14-SEP-1999;   99US-0153872.
PR   19-NOV-1999;   99US-0166573.
XX
PA   (CVTH-) CV THERAPEUTICS INC.
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Lawn RM,  Wade D,  Oram JF,  Garvin M;
XX
DR   WPI; 2001-137811/14.
DR   P-PSDB; AAB31366.
XX
PT   Adenosine triphosphate (ATP) binding cassette protein (ABC) 1
PT   polynucleotides and polypeptides, useful for treatment of heart disease
PT   and other disorders associated with hypercholesterolemia and
PT   atherosclerosis -
XX
PS   Claim 27; Page 144-150; 211pp; English.
XX
CC   The present sequence encodes a human adenosine triphosphate (ATP)
CC   binding cassette protein (ABC) 1 polypeptide, and is isolated from
CC   a Tangier disease patient. ABC1 resides in cell membranes and utilises
CC   ATP hydrolysis to transport a wide variety of substrates across the
CC   plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated
CC   mobilisation of intracellular cholesterol stores. ABC1 is defective in
CC   Tangier disease, a genetic disorder characterised by abnormal
CC   HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome
CC   9q22-9q31. The ABC1 genes and proteins are useful for developing
CC   pharmaceutical agents for the treatment of heart disease and other
CC   disorders associated with hypercholesterolemia and atherosclerosis. The
CC   genes are useful for developing screening assays to screen for compounds
CC   that regulate the expression of genes associated with cholesterol
CC   transport. The genes and proteins are also useful for are also useful
CC   as diagnostic indicators of cardiovascular disease and other disorders
CC   associated with hypercholesterolemia.
XX
SQ   Sequence 10474 BP; 2906 A; 2305 C; 2416 G; 2843 T; 4 other;


 Query Match            48.4%;  Score 77;  DB 22;  Length 10474;
 Best Local Similarity  99.2%;  Pred. No. 5e-29;

  Matches  127;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     32 agctctggccgctgccttccagggctcccgagccacacgctgggcgtgctggctgaggga 91
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    261 agctctggccgctgccttccagggctcccgagccacacgctgggcgtgctggctgaggga 320

Qy     92 acatggcatgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaa 151
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    321 acatggcttgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaa 380

Qy    152 gacaaaca 159
          ||||||||
Db    381 gacaaaca 388


RESULT   6
AAF24708
ID   AAF24708 standard; DNA; 10474 BP.
XX
AC   AAF24708;
XX
DT   20-APR-2001  (first entry)
XX
DE   Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.
XX
KW   Human; adenosine triphosphate binding cassette protein 1; ABC1;
KW   apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW   chromosome 9q22-9q31; heart disease; hypercholesterolemia;
KW   atherosclerosis; cholesterol transport; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             323..7108
FT                   /*tag=  a
FT                   /product= "defective ABC1 polypeptide"
XX
PN   WO200078971-A2.
XX
PD   28-DEC-2000.
XX
PF   16-JUN-2000; 2000WO-US16591.
XX
PR   18-JUN-1999;   99US-0140264.
PR   14-SEP-1999;   99US-0153872.
PR   19-NOV-1999;   99US-0166573.
XX
PA   (CVTH-) CV THERAPEUTICS INC.
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Lawn RM,  Wade D,  Oram JF,  Garvin M;
XX
DR   WPI; 2001-137811/14.
DR   P-PSDB; AAB31367.
XX
PT   Adenosine triphosphate (ATP) binding cassette protein (ABC) 1
PT   polynucleotides and polypeptides, useful for treatment of heart disease
PT   and other disorders associated with hypercholesterolemia and
PT   atherosclerosis -
XX
PS   Claim 30; Page 165-172; 211pp; English.
XX
CC   The present sequence encodes a human adenosine triphosphate (ATP)
CC   binding cassette protein (ABC) 1 polypeptide, and is isolated from
CC   a Tangier disease patient. ABC1 resides in cell membranes and utilises
CC   ATP hydrolysis to transport a wide variety of substrates across the
CC   plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated
CC   mobilisation of intracellular cholesterol stores. ABC1 is defective in
CC   Tangier disease, a genetic disorder characterised by abnormal
CC   HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome
CC   9q22-9q31. The ABC1 genes and proteins are useful for developing
CC   pharmaceutical agents for the treatment of heart disease and other
CC   disorders associated with hypercholesterolemia and atherosclerosis. The
```

```
CC   genes are useful for developing screening assays to screen for compounds
CC   that regulate the expression of genes associated with cholesterol
CC   transport. The genes and proteins are also useful for are also useful
CC   as diagnostic indicators of cardiovascular disease and other disorders
CC   associated with hypercholesterolemia.
XX
SQ   Sequence 10474 BP; 2907 A; 2304 C; 2415 G; 2844 T; 4 other;


  Query Match            48.4%;  Score 77;  DB 22;  Length 10474;
  Best Local Similarity  99.2%;  Pred. No. 5e-29;
  Matches  127;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     32 agctctggccgctgccttccagggctcccgagccacacgctgggcgtgctggctgaggga 91
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    261 agctctggccgctgccttccagggctcccgagccacacgctgggcgtgctggctgaggga 320

Qy     92 acatggcatgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaa 151
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    321 acatggcttgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaa 380

Qy    152 gacaaaca 159
          ||||||||
Db    381 gacaaaca 388


RESULT   7
AAS04035
ID   AAS04035 standard; cDNA; 446 BP.
XX
AC   AAS04035;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Partial human ABC1 cDNA sequence.
XX
KW   Human; ABC1 gene; atherosclerosis; reverse transport; cholesterol;
KW   cardiovascular; neurological; Tangier disease;, LCAT deficiency;
KW   lecithin-cholesterol acetyltransferase; malaria; diabetes; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             185..438
FT                   /*tag=  a
FT                   /product= "Human ABC1 protein, amino acids 1 to 60"
XX
PN   WO200130848-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-EP10886.
XX
PR   26-OCT-1999;   99EP-0402668.
PR   01-MAR-2000; 2000US-0186260.
XX
PA   (AVET ) AVENTIS PHARMA SA.
XX
PI   Denefle P,  Rosier-Montus M,  Arnould-Reguigne I,  Prades C,  Naudin L;
PI   Lemoine C,  Duverger N,  Jaye M,  Searfoss GH,  Remaley A,  Brewer HB;
PI   Dean M;
XX
DR   WPI; 2001-316327/33.
DR   P-PSDB; AAU02176.
XX
PT   New human ABC1 nucleic acids and polypeptides for treating
PT   atherosclerosis, malaria and diabetes -
XX
PS   Example 2; Page 167; 368pp; English.
XX
CC   The sequence represents the partial coding sequence of human ABC1,
CC   which encodes amino acids 1-60 of the human ABC1 protein. The nucleic
CC   acid sequence, primers and probes derived from the ABC1 sequence, and
CC   polypeptides and vectors are useful for the prevention of
CC   atherosclerosis, in a subject affected by a dysfunction in the reverse
CC   transport of cholesterol. The polypeptide encoded by the ABC1 gene is
CC   useful for screening for an active ingredient for the prevention or
CC   treatment of a disease resulting from dysfunction in the reverse
CC   transport of cholesterol. The nucleic acids and polypeptides are also
CC   useful for treating and preventing cardiovascular and neurological
CC   pathologies, and other diseases e.g. Tangier disease, lecithin-
CC   cholesterol (LCAT) deficiency, malaria and diabetes.
XX
SQ   Sequence 446 BP; 96 A; 123 C; 112 G; 115 T; 0 other;


  Query Match            37.7%;  Score 60;  DB 22;  Length 446;
  Best Local Similarity  100.0%;  Pred. No. 1.6e-20;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    191 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 250


RESULT   8
ABA09200
ID   ABA09200 standard; cDNA; 7086 BP.
XX
AC   ABA09200;
XX
DT   11-JAN-2002  (first entry)
XX
DE   Human ABCA1 homologue-encoding cDNA, SEQ ID NO:976.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW   haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW   inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW   proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW   myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW   chronic inflammatory condition; proliferative retinopathy;
KW   atherosclerosis; coronary heart disease; arterial ischaemia;
KW   bone disorder; osteoporosis; vascular growth disorder;
KW   tissue regeneration; wound healing; infection; immune disorder;
KW   cell culture; drug screening; gene therapy; antiinflammatory;
KW   antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW   cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW   antifungal; vulnerary; antiulcer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157188-A2.
XX
PD   09-AUG-2001.
XX
PF   05-FEB-2001; 2001WO-US03800.
XX
PR   03-FEB-2000; 2000US-0496914.
PR   27-APR-2000; 2000US-0560875.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-457740/49.
DR   P-PSDB; ABB11956.
XX
PT   Human proteins and DNA encoding sequences useful for preventing,
PT   treating or ameliorating a medical condition in a mammalian subject
PT   e.g. arthritis and cancer -
XX
PS   Claim 1; Page 833-835; 1963pp; English.
XX
CC   Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC   sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC   invention also relates to vectors and recombinant host cells comprising a
```

CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g.,myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a cDNA encoding a
CC novel human polypeptide of the invention.
XX
SQ Sequence 7086 BP; 1773 A; 1739 C; 1859 G; 1715 T; 0 other;

Query Match 37.7%; Score 60; DB 22; Length 7086;
Best Local Similarity 100.0%; Pred. No. 1.6e-20;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 310 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 369

RESULT 9
AAK52667
ID AAK52667 standard; cDNA; 7086 BP.
XX
AC AAK52667;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 2196.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US04098.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR WPI; 2001-476283/51.
DR P-PSDB; AAM79534.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
PS Claim 1; Page 4558-4560; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
SQ Sequence 7086 BP; 1773 A; 1739 C; 1859 G; 1715 T; 0 other;

Query Match 37.7%; Score 60; DB 22; Length 7086;
Best Local Similarity 100.0%; Pred. No. 1.6e-20;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 310 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 369

RESULT 10
AAD21326
ID AAD21326 standard; DNA; 7260 BP.
XX
AC AAD21326;
XX
DT 28-JAN-2002 (first entry)
XX
DE Human ATP binding cassette transporter 1 (ABC1) gene.
XX
KW Human; ATP binding cassette transporter 1; ABC1; coronary heart disease;
KW dermatological; atherosclerosis; cardiovascular; inflammatory disease;
KW psoriasis; lipid disorder; antibacterial; septic shock; gene therapy;
KW immunosuppressive; lupus erythematosus; rheumatoid arthritis; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 321..7106
FT /*tag= a
FT /product= "Human ABC1 protein"
XX
PN EP1136552-A1.
XX
PD 26-SEP-2001.
XX

PF 20-MAR-2000; 2000EP-0105820.
XX
PR 20-MAR-2000; 2000EP-0105820.
XX
PA (FARB) BAYER AG.
XX
PI Schmitz G, Bodzioch M;
XX
DR WPI; 2001-640388/74.
DR P-PSDB; AAE13022.
XX
PT New adenosine triphosphate binding cassette transporter-1 gene
PT polymorphisms, useful for diagnosing and treating lipid disorders,
PT cardiovascular diseases and inflammatory diseases -
XX
PS Example 1; Fig 1; 48pp; English.
XX
CC The invention relates to four common polymorphisms in the gene encoding
CC ATP-binding cassette transporter-1 (ABC1). ABC1 is associated with
CC decreased ApoA-1 mediated efflux of cholesterol. The polymorphisms in
CC ABC1 directly affects cellular lipid homeostasis, which is a key factor
CC in the atherogenetic processes. The ABC1 polymorphisms are useful for
CC diagnosing and treating lipid disorders, cardiovascular diseases
CC (coronary heart disease, atherosclerosis) and inflammatory diseases
CC (psoriasis, lupus erythematosus). The identification of ABC1 as a
CC transporter for interleukin-1beta (IL-1beta) identifies this gene as
CC a candidate for treatment of inflammatory diseases including rheumatoid
CC arthritis and septic shock. The present sequence is human ABC1 gene.
XX
SQ Sequence 7260 BP; 1834 A; 1765 C; 1905 G; 1756 T; 0 other;

Query Match 37.7%; Score 60; DB 22; Length 7260;
Best Local Similarity 100.0%; Pred. No. 1.6e-20;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 327 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 386

RESULT 11
AAI70315
ID AAI70315 standard; cDNA; 7260 BP.
XX
AC AAI70315;
XX
DT 07-JAN-2002 (first entry)
XX
DE Human ATP binding cassette transporter 1 (ABC1) cDNA.
XX
KW ATP binding cassette transporter 1; ABC1; human; lipid disorder;
KW cholesterol; cardiovascular disease; inflammatory disease;
KW antiinflammatory; antilipaemic; antipsoriatic; dermatological;
KW Tangier disease; coronary heart disease; diagnosis; gene therapy;
KW polymorphism; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 321..7106
FT /*tag= a
FT CDS 501..7106
FT /*tag= b
FT /note= "alternative open reading frame of AAI70314"
FT variation replace(976,A)
FT /*tag= c
FT variation replace(1516,C)
FT /*tag= d
FT variation replace(2969,G)
FT /*tag= e
FT variation replace(3836,C)
FT /*tag= f
XX
PN EP1136554-A1.
XX
PD 26-SEP-2001.
XX
PF 24-MAR-2000; 2000EP-0106401.
XX
PR 24-MAR-2000; 2000EP-0106401.
XX
PA (FARB) BAYER AG.
XX
PI Schmitz G, Bodzioch M;
XX
DR WPI; 2001-640389/74.
DR P-PSDB; AAM50228.
XX
PT New adenosine triphosphate binding cassette transporter gene
PT polymorphisms, useful for diagnosing and treating lipid disorders,
PT cardiovascular diseases and inflammatory diseases -
XX
PS Disclosure; Page 26-28; 41pp; English.
XX
CC The present sequence is that of cDNA encoding the human adenosine
CC triphosphate (ATP) binding cassette transporter 1 (ABC1) protein
CC (see AAM50227). The sequence includes an extended open reading
CC frame (ORF) to that provided by the sequence in AAI70314, using
CC an alternative ATG codon as initiation codon and thereby adding an
CC extra 40 N-terminal amino acids to the encoded ABC1 protein (see
CC AAM50228). The invention provides 4 common polymorphisms in the
CC ABC1 gene. These were identified by sequencing the ABC1 gene in
CC different Tangier kindreds. In the variant genes (numbering as in
CC AAI70314), G is changed to A at position 596, T is changed to C at
CC position 1136, A is changed to G at position 2589 or G is changed
CC to C at position 3456, or any combination of these. All of these
CC polymorphisms alter the amino acid sequence of ABC1 and therefore
CC may affect its function. The 2 most common polymorphisms (G596A)
CC and A2589G) are both associated with a decreased in vitro ApoA-I
CC mediated efflux of cholesterol from mononuclear phagocytes, a
CC feature typical of Tangier disease. 3 Of the variants (G596A,
CC A2589G and G3456C) are significantly increased in a population of
CC men having low high density lipoprotein-cholesterol levels and
CC established coronary heart disease (CHD) relative to CHD-free
CC control subjects. The use of the provided ABC1 polymorphisms for
CC the diagnosis and treatment of lipid disorders, cardiovascular
CC diseases, and inflammatory diseases (e.g. psoriasis, lupus
CC erythematodes) is claimed. Modulation of ABC1 transcripts or
CC proteins by antisense or ribozyme technology or RNA decoys is also
CC claimed.
XX
SQ Sequence 7260 BP; 1834 A; 1765 C; 1905 G; 1756 T; 0 other;

Query Match 37.7%; Score 60; DB 22; Length 7260;
Best Local Similarity 100.0%; Pred. No. 1.6e-20;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 327 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 386

RESULT 12
AAK51683
ID AAK51683 standard; cDNA; 7281 BP.
XX
AC AAK51683;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 228.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US04098.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR WPI; 2001-476283/51.
DR P-PSDB; AAM78550.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
PS Claim 1; Page 1086-1096; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
SQ Sequence 7281 BP; 1831 A; 1773 C; 1915 G; 1762 T; 0 other;

Query Match 37.7%; Score 60; DB 22; Length 7281;
Best Local Similarity 100.0%; Pred. No. 1.6e-20;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 348 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 407

RESULT 13
AAC69388
ID AAC69388 standard; cDNA; 7857 BP.
XX
AC AAC69388;
XX
DT 29-JAN-2001 (first entry)
XX
DE Human ABC1 cholesterol transporter FHA-3 mutant cDNA (delta 5752-5757).
XX
KW Human ABC1 cholesterol transporter; chromosome 9q31;
KW ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW cardiovascular disease; coronary artery disease; coronary restenosis;
KW cerebrovascular disease; peripheral vascular disease;
KW Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW prognosis; prophylaxis; drug screening; transgenic animal; mutant; ss.
XX
OS Homo sapiens.
XX
PN WO200055318-A2.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-IB00532.
XX
PR 15-MAR-1999; 99US-0124702.
PR 08-JUN-1999; 99US-0138048.
PR 17-JUN-1999; 99US-0139600.
PR 01-SEP-1999; 99US-0151977.
XX
PA (UYBR-) UNIV BRITISH COLUMBIA.
PA (XENO-) XENON BIORESEARCH INC.
XX
PI Hayden MR, Wilson AR, Pimstone SN;
XX
DR WPI; 2000-587528/55.
DR P-PSDB; AAB38107.
XX
PT New ABC1 polypeptide is useful for treating diseases associated with
PT ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
PT disease and cancer -
XX
PS Examples; Page -; 229pp; English.
XX
CC The invention relates to the human ABC1 cholesterol transporter protein
CC (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC a member of the ATP-binding cassette (ABC transporter) superfamily of
CC proteins, and plays a crucial role in cholesterol transport, particularly
CC intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC located on chromosome 9q31, and mutations in this gene are associated
CC with two genetic HDL (high density lipoprotein) deficiency disorders,
CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC are distinguishable in that TD is an autosomal recessive disorder, while
CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC cholesterol") in the blood correlate with a high risk of cardiovascular
CC disease, particularly coronary artery disease, but also cerebrovascular
CC disease, coronary restenosis, and peripheral vascular disease.
CC Conversely, a high level of HDL has protective effects against
CC cardiovascular disease. The invention provides genetic constructs and
CC transgenic cells and non-human animals comprising human ABC1 nucleic
CC acids, and methods of gene therapy for the treatment or prevention of
CC cardiovascular disease comprising the administration of an expression
CC vector encoding ABC1 or an active fragment thereof. The invention also
CC encompasses compounds which mimic ABC1 activity, compounds which
CC stimulate ABC1 expression and methods of screening for such compounds.
CC It further relates to methods for determining whether a patient has an
CC increased risk for cardiovascular disease due to polymorphisms in the
CC ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
CC or prevent cardiovascular disease, especially coronary artery disease,
CC cerebrovascular disease, coronary restenosis or peripheral vascular
CC disease. They may also be used in the treatment of diseases associated
CC with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC The invention specifically excludes proteins with the exact amino acid
CC sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
CC acid with the exact sequence as GenBank Accession No: AJ012376.1. The
CC present sequence represents cDNA encoding a mutant human ABC1 cholesterol
CC transporter associated with an altered cholesterol level and therefore an
CC altered risk of cardiovascular disease.
CC Note: The present sequence is not shown in the specification, but is
CC derived from the native human ABC1 cDNA shown on pages 157-160.
XX

```
SQ   Sequence 7857 BP; 2011 A; 1860 C; 2008 G; 1977 T; 1 other;

  Query Match             37.7%;  Score 60;  DB 21;  Length 7857;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-20;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      81 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 140


RESULT  14
AAF83826
ID   AAF83826 standard; DNA; 7860 BP.
XX
AC   AAF83826;
XX
DT   06-AUG-2001  (first entry)
XX
DE   Human ABC1 nucleotide sequence.
XX
KW   ABC1; antilipemic; cholesterol; inhibitor; low density lipoprotein;
KW   LDL; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             75..3341
FT                   /*tag=  a
FT                   /product= "partial ABC1 protein"
FT                   /note= "the coding sequence continues beyond nucleotide
FT                          3341, possibly till position 6860 as identified
FT                          by translating the present sequence; part of the
FT                          corresponding protein is missing and nucleotide
FT                          3341 corresponds to the last amino acid residue
FT                          (position 1089) as indicated in the
FT                           specification"
XX
PN   WO200132184-A2.
XX
PD   10-MAY-2001.
XX
PF   01-NOV-2000; 2000WO-US30109.
XX
PR   01-NOV-1999;   99US-0162803.
PR   30-JUN-2000; 2000US-0215564.
XX
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Attie AD,  Cook M,  Gray-Keller MP,  Hayden MR,  Pimstone S;
PI   Brooks-Wilson A;
XX
DR   WPI; 2001-335779/35.
DR   P-PSDB; AAB62691.
XX
PT   New method for inhibiting cholesterol uptake in the gut comprises
PT   administration of an inhibitor of an ABC1 protein  -
XX
PS   Disclosure; Page 34-36; 41pp; English.
XX
CC   The invention relates to a new method for inhibiting cholesterol uptake
CC   in the gut that comprises administration of an inhibitor of an ABC1
CC   protein. The method is useful for: lowering levels of LDL (low density
CC   lipoprotein) cholesterol by reducing the activity of ABC1 protein in the
CC   intestinal cells and the abundance of the ABC1 protein in the individual
CC   by inhibiting the activity of the protein; identifying drugs that can
CC   lower serum cholesterol and LDL levels comprises assaying the drug to
CC   test if it can bind to an ABC1 protein; testing LDL cholesterol lowering
CC   agents; and for modulation of ABC1 biological activity. The present
CC   sequence represents a human ABC1 nucleotide sequence.
XX
SQ   Sequence 7860 BP; 2013 A; 1861 C; 2009 G; 1977 T; 0 other;

  Query Match             37.7%;  Score 60;  DB 22;  Length 7860;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-20;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      81 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 140


RESULT  15
AAF92835
ID   AAF92835 standard; DNA; 7860 BP.
XX
AC   AAF92835;
XX
DT   17-MAY-2001  (first entry)
XX
DE   Human ABC1 cDNA.
XX
KW   High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ss.
XX
OS   Homo sapiens.
XX
PN   WO200115676-A2.
XX
PD   08-MAR-2001.
XX
PF   01-SEP-2000; 2000WO-IB01492.
XX
PR   01-SEP-1999;   99US-0151977.
PR   15-MAR-2000; 2000US-0526193.
PR   23-JUN-2000; 2000US-0213958.
XX
PA   (UYBR-) UNIV BRITISH COLUMBIA.
PA   (XENO-) XENON GENETICS INC.
XX
PI   Hayden MR,  Brooks-Wilson AR,  Pimstone SN,  Clee SM;
XX
DR   WPI; 2001-244356/25.
XX
PT   Treating a lower than normal high density lipoprotein-cholesterol
PT   (HDL-C) level, a higher than normal triglyceride level, or a
PT   cardiovascular disease, by administering a compound that modulates LXR-
PT   or RXR-mediated transcriptional activity -
XX
PS   Disclosure; Fig 2; 317pp; English.
XX
CC   The present invention relates to a method for treating a patient
CC   diagnosed as having a lower than normal high density
CC   lipoprotein-cholesterol (HDL-C) level, a higher than normal
CC   triglyceride level, or a cardiovascular disease, involving
CC   administering a compound that modulates LXR- or RXR-mediated
CC   transcriptional activity or ABC1 expression or activity.
CC   The LXR gene product may be used in an assay to identify
CC   compounds useful for the treatment of a disease or condition selected a
CC   lower than normal HDL cholesterol level, a higher than normal
CC   triglyceride level, and a cardiovascular disease.
XX
SQ   Sequence 7860 BP; 2014 A; 1860 C; 2008 G; 1978 T; 0 other;

  Query Match             37.7%;  Score 60;  DB 22;  Length 7860;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-20;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      81 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 140


RESULT  16
```

```
AAC69387
ID   AAC69387 standard; cDNA; 7861 BP.
XX
AC   AAC69387;
XX
DT   29-JAN-2001  (first entry)
XX
DE   Human ABC1 cholesterol transporter FHA-1 mutant cDNA (delta 2151-2153).
XX
KW   Human ABC1 cholesterol transporter; chromosome 9q31;
KW   ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW   Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW   cardiovascular disease; coronary artery disease; coronary restenosis;
KW   cerebrovascular disease; peripheral vascular disease;
KW   Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW   X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW   prognosis; prophylaxis; drug screening; transgenic animal; mutant; ss.
XX
OS   Homo sapiens.
XX
PN   WO200055318-A2.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-IB00532.
XX
PR   15-MAR-1999;   99US-0124702.
PR   08-JUN-1999;   99US-0138048.
PR   17-JUN-1999;   99US-0139600.
PR   01-SEP-1999;   99US-0151977.
XX
PA   (UYBR-) UNIV BRITISH COLUMBIA.
PA   (XENO-) XENON BIORESEARCH INC.
XX
PI   Hayden MR,  Wilson AR,  Pimstone SN;
XX
DR   WPI; 2000-587528/55.
DR   P-PSDB; AAB38106.
XX
PT   New ABC1 polypeptide is useful for treating diseases associated with
PT   ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
PT   disease and cancer -
XX
PS   Examples; Page -; 229pp; English.
XX
CC   The invention relates to the human ABC1 cholesterol transporter protein
CC   (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC   a member of the ATP-binding cassette (ABC transporter) superfamily of
CC   proteins, and plays a crucial role in cholesterol transport, particularly
CC   intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC   involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC   located on chromosome 9q31, and mutations in this gene are associated
CC   with two genetic HDL (high density lipoprotein) deficiency disorders,
CC   Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC   are distinguishable in that TD is an autosomal recessive disorder, while
CC   FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC   cholesterol") in the blood correlate with a high risk of cardiovascular
CC   disease, particularly coronary artery disease, but also cerebrovascular
CC   disease, coronary restenosis, and peripheral vascular disease.
CC   Conversely, a high level of HDL has protective effects against
CC   cardiovascular disease. The invention provides genetic constructs and
CC   transgenic cells and non-human animals comprising human ABC1 nucleic
CC   acids, and methods of gene therapy for the treatment or prevention of
CC   cardiovascular disease comprising the administration of an expression
CC   vector encoding ABC1 or an active fragment thereof. The invention also
CC   encompasses compounds which mimic ABC1 activity, compounds which
CC   stimulate ABC1 expression and methods of screening for such compounds.
CC   It further relates to methods for determining whether a patient has an
CC   increased risk for cardiovascular disease due to polymorphisms in the
CC   ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
CC   or prevent cardiovascular disease, especially coronary artery disease,
CC   cerebrovascular disease, coronary restenosis or peripheral vascular
CC   disease. They may also be used in the treatment of diseases associated
CC   with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC   disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC   The invention specifically excludes proteins with the exact amino acid
CC   sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
CC   acid with the exact sequence as GenBank Accession No: AJ012376.1. The
CC   present sequence represents cDNA encoding a mutant human ABC1 cholesterol
CC   transporter associated with an altered cholesterol level and therefore an
CC   altered risk of cardiovascular disease.
CC   Note: The present sequence is not shown in the specification, but is
CC   derived from the native human ABC1 cDNA shown on pages 157-160.
XX
SQ   Sequence 7861 BP; 2014 A; 1859 C; 2011 G; 1976 T; 1 other;


  Query Match             37.7%;  Score 60;  DB 21;  Length 7861;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-20;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 140


RESULT  17
AAC69120
ID   AAC69120 standard; cDNA; 7864 BP.
XX
AC   AAC69120;
XX
DT   29-JAN-2001  (first entry)
XX
DE   Human ABC1 cholesterol transporter cDNA.
XX
KW   Human ABC1 cholesterol transporter; chromosome 9q31;
KW   ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW   Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW   cardiovascular disease; coronary artery disease; coronary restenosis;
KW   cerebrovascular disease; peripheral vascular disease;
KW   Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW   X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW   prognosis; prophylaxis; drug screening; transgenic animal; ss.
XX
OS   Homo sapiens.
XX
PN   WO200055318-A2.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-IB00532.
XX
PR   15-MAR-1999;   99US-0124702.
PR   08-JUN-1999;   99US-0138048.
PR   17-JUN-1999;   99US-0139600.
PR   01-SEP-1999;   99US-0151977.
XX
PA   (UYBR-) UNIV BRITISH COLUMBIA.
PA   (XENO-) XENON BIORESEARCH INC.
XX
PI   Hayden MR,  Wilson AR,  Pimstone SN;
XX
DR   WPI; 2000-587528/55.
DR   P-PSDB; AAB38082.
XX
PT   New ABC1 polypeptide is useful for treating diseases associated with
PT   ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
PT   disease and cancer -
XX
PS   Claim 13; Page 157-160; 229pp; English.
XX
CC   The invention relates to the human ABC1 cholesterol transporter protein
CC   (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC   a member of the ATP-binding cassette (ABC transporter) superfamily of
CC   proteins, and plays a crucial role in cholesterol transport, particularly
```

CC intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC located on chromosome 9q31, and mutations in this gene are associated
CC with two genetic HDL (high density lipoprotein) deficiency disorders,
CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC are distinguishable in that TD is an autosomal recessive disorder, while
CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC cholesterol") in the blood correlate with a high risk of cardiovascular
CC disease, particularly coronary artery disease, but also cerebrovascular
CC disease, coronary restenosis, and peripheral vascular disease.
CC Conversely, a high level of HDL has protective effects against
CC cardiovascular disease. The invention provides genetic constructs and
CC transgenic cells and non-human animals comprising human ABC1 nucleic
CC acids, and methods of gene therapy for the treatment or prevention of
CC cardiovascular disease comprising the administration of an expression
CC vector encoding ABC1 or an active fragment thereof. The invention also
CC encompasses compounds which mimic ABC1 activity, compounds which
CC stimulate ABC1 expression and methods of screening for such compounds.
CC It further relates to methods for determining whether a patient has an
CC increased risk for cardiovascular disease due to polymorphisms in the
CC ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
CC or prevent cardiovascular disease, especially coronary artery disease,
CC cerebrovascular disease, coronary restenosis or peripheral vascular
CC disease. They may also be used in the treatment of diseases associated
CC with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC The invention specifically excludes proteins with the exact amino acid
CC sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
CC acid with the exact sequence as GenBank Accession No: AJ012376.1. The
CC present sequence represents cDNA encoding the human ABC1 cholesterol
CC transporter.
XX
SQ Sequence 7864 BP; 2014 A; 1860 C; 2011 G; 1978 T; 1 other;

Query Match 37.7%; Score 60; DB 21; Length 7864;
Best Local Similarity 100.0%; Pred. No. 1.6e-20;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 81 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 140

RESULT 18
AAC69385
ID AAC69385 standard; cDNA; 7864 BP.
XX
AC AAC69385;
XX
DT 29-JAN-2001 (first entry)
XX
DE Human ABC1 cholesterol transporter TD-1 mutant cDNA (T4503C).
XX
KW Human ABC1 cholesterol transporter; chromosome 9q31;
KW ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW cardiovascular disease; coronary artery disease; coronary restenosis;
KW cerebrovascular disease; peripheral vascular disease;
KW Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW prognosis; prophylaxis; drug screening; transgenic animal; mutant; ss.
XX
OS Homo sapiens.
XX
PN WO200055318-A2.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-IB00532.
XX
PR 15-MAR-1999; 99US-0124702.
PR 08-JUN-1999; 99US-0138048.
PR 17-JUN-1999; 99US-0139600.
PR 01-SEP-1999; 99US-0151977.
XX
PA (UYBR-) UNIV BRITISH COLUMBIA.
PA (XENO-) XENON BIORESEARCH INC.
XX
PI Hayden MR, Wilson AR, Pimstone SN;
XX
DR WPI; 2000-587528/55.
DR P-PSDB; AAB38104.
XX
PT New ABC1 polypeptide is useful for treating diseases associated with
PT ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
PT disease and cancer -
XX
PS Examples; Page -; 229pp; English.
XX
CC The invention relates to the human ABC1 cholesterol transporter protein
CC (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC a member of the ATP-binding cassette (ABC transporter) superfamily of
CC proteins, and plays a crucial role in cholesterol transport, particularly
CC intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC located on chromosome 9q31, and mutations in this gene are associated
CC with two genetic HDL (high density lipoprotein) deficiency disorders,
CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC are distinguishable in that TD is an autosomal recessive disorder, while
CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC cholesterol") in the blood correlate with a high risk of cardiovascular
CC disease, particularly coronary artery disease, but also cerebrovascular
CC disease, coronary restenosis, and peripheral vascular disease.
CC Conversely, a high level of HDL has protective effects against
CC cardiovascular disease. The invention provides genetic constructs and
CC transgenic cells and non-human animals comprising human ABC1 nucleic
CC acids, and methods of gene therapy for the treatment or prevention of
CC cardiovascular disease comprising the administration of an expression
CC vector encoding ABC1 or an active fragment thereof. The invention also
CC encompasses compounds which mimic ABC1 activity, compounds which
CC stimulate ABC1 expression and methods of screening for such compounds.
CC It further relates to methods for determining whether a patient has an
CC increased risk for cardiovascular disease due to polymorphisms in the
CC ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
CC or prevent cardiovascular disease, especially coronary artery disease,
CC cerebrovascular disease, coronary restenosis or peripheral vascular
CC disease. They may also be used in the treatment of diseases associated
CC with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC The invention specifically excludes proteins with the exact amino acid
CC sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
CC acid with the exact sequence as GenBank Accession No: AJ012376.1. The
CC present sequence represents cDNA encoding a mutant human ABC1 cholesterol
CC transporter associated with an altered cholesterol level and therefore an
CC altered risk of cardiovascular disease.
CC Note: The present sequence is not shown in the specification, but is
CC derived from the native human ABC1 cDNA shown on pages 157-160.
XX
SQ Sequence 7864 BP; 2014 A; 1861 C; 2011 G; 1977 T; 1 other;

Query Match 37.7%; Score 60; DB 21; Length 7864;
Best Local Similarity 100.0%; Pred. No. 1.6e-20;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 81 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 140

RESULT 19
AAC69386
ID AAC69386 standard; cDNA; 7864 BP.
XX
AC AAC69386;

XX
DT 29-JAN-2001 (first entry)
XX
DE Human ABC1 cholesterol transporter TD-2 mutant cDNA (A1864G).
XX
KW Human ABC1 cholesterol transporter; chromosome 9q31;
KW ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW cardiovascular disease; coronary artery disease; coronary restenosis;
KW cerebrovascular disease; peripheral vascular disease;
KW Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW prognosis; prophylaxis; drug screening; transgenic animal; mutant; ss.
XX
OS Homo sapiens.
XX
PN WO200055318-A2.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-IB00532.
XX
PR 15-MAR-1999; 99US-0124702.
PR 08-JUN-1999; 99US-0138048.
PR 17-JUN-1999; 99US-0139600.
PR 01-SEP-1999; 99US-0151977.
XX
PA (UYBR-) UNIV BRITISH COLUMBIA.
PA (XENO-) XENON BIORESEARCH INC.
XX
PI Hayden MR, Wilson AR, Pimstone SN;
XX
DR WPI; 2000-587528/55.
DR P-PSDB; AAB38105.
XX
PT New ABC1 polypeptide is useful for treating diseases associated with
PT ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
PT disease and cancer -
XX
PS Examples; Page -; 229pp; English.
XX
CC The invention relates to the human ABC1 cholesterol transporter protein
CC (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC a member of the ATP-binding cassette (ABC transporter) superfamily of
CC proteins, and plays a crucial role in cholesterol transport, particularly
CC intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC located on chromosome 9q31, and mutations in this gene are associated
CC with two genetic HDL (high density lipoprotein) deficiency disorders,
CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC are distinguishable in that TD is an autosomal recessive disorder, while
CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC cholesterol") in the blood correlate with a high risk of cardiovascular
CC disease, particularly coronary artery disease, but also cerebrovascular
CC disease, coronary restenosis, and peripheral vascular disease.
CC Conversely, a high level of HDL has protective effects against
CC cardiovascular disease. The invention provides genetic constructs and
CC transgenic cells and non-human animals comprising human ABC1 nucleic
CC acids, and methods of gene therapy for the treatment or prevention of
CC cardiovascular disease comprising the administration of an expression
CC vector encoding ABC1 or an active fragment thereof. The invention also
CC encompasses compounds which mimic ABC1 activity, compounds which
CC stimulate ABC1 expression and methods of screening for such compounds.
CC It further relates to methods for determining whether a patient has an
CC increased risk for cardiovascular disease due to polymorphisms in the
CC ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
CC or prevent cardiovascular disease, especially coronary artery disease,
CC cerebrovascular disease, coronary restenosis or peripheral vascular
CC disease. They may also be used in the treatment of diseases associated
CC with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC The invention specifically excludes proteins with the exact amino acid
CC sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
CC acid with the exact sequence as GenBank Accession No: AJ012376.1. The
CC present sequence represents cDNA encoding a mutant human ABC1 cholesterol
CC transporter associated with an altered cholesterol level and therefore an
CC altered risk of cardiovascular disease.
CC Note: The present sequence is not shown in the specification, but is
CC derived from the native human ABC1 cDNA shown on pages 157-160.
XX
SQ Sequence 7864 BP; 2013 A; 1860 C; 2012 G; 1978 T; 1 other;

Query Match 37.7%; Score 60; DB 21; Length 7864;
Best Local Similarity 100.0%; Pred. No. 1.6e-20;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 81 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 140

RESULT 20
AAC69389
ID AAC69389 standard; cDNA; 7864 BP.
XX
AC AAC69389;
XX
DT 29-JAN-2001 (first entry)
XX
DE Human ABC1 cholesterol transporter FHA-2 mutant cDNA (C6504T).
XX
KW Human ABC1 cholesterol transporter; chromosome 9q31;
KW ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW cardiovascular disease; coronary artery disease; coronary restenosis;
KW cerebrovascular disease; peripheral vascular disease;
KW Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW prognosis; prophylaxis; drug screening; transgenic animal; mutant; ss.
XX
OS Homo sapiens.
XX
PN WO200055318-A2.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-IB00532.
XX
PR 15-MAR-1999; 99US-0124702.
PR 08-JUN-1999; 99US-0138048.
PR 17-JUN-1999; 99US-0139600.
PR 01-SEP-1999; 99US-0151977.
XX
PA (UYBR-) UNIV BRITISH COLUMBIA.
PA (XENO-) XENON BIORESEARCH INC.
XX
PI Hayden MR, Wilson AR, Pimstone SN;
XX
DR WPI; 2000-587528/55.
DR P-PSDB; AAB38108.
XX
PT New ABC1 polypeptide is useful for treating diseases associated with
PT ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
PT disease and cancer -
XX
PS Examples; Page -; 229pp; English.
XX
CC The invention relates to the human ABC1 cholesterol transporter protein
CC (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC a member of the ATP-binding cassette (ABC transporter) superfamily of
CC proteins, and plays a crucial role in cholesterol transport, particularly
CC intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC located on chromosome 9q31, and mutations in this gene are associated
CC with two genetic HDL (high density lipoprotein) deficiency disorders,

```
CC   Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC   are distinguishable in that TD is an autosomal recessive disorder, while
CC   FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC   cholesterol") in the blood correlate with a high risk of cardiovascular
CC   disease, particularly coronary artery disease, but also cerebrovascular
CC   disease, coronary restenosis, and peripheral vascular disease.
CC   Conversely, a high level of HDL has protective effects against
CC   cardiovascular disease. The invention provides genetic constructs and
CC   transgenic cells and non-human animals comprising human ABC1 nucleic
CC   acids, and methods of gene therapy for the treatment or prevention of
CC   cardiovascular disease comprising the administration of an expression
CC   vector encoding ABC1 or an active fragment thereof. The invention also
CC   encompasses compounds which mimic ABC1 activity, compounds which
CC   stimulate ABC1 expression and methods of screening for such compounds.
CC   It further relates to methods for determining whether a patient has an
CC   increased risk for cardiovascular disease due to polymorphisms in the
CC   ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
CC   or prevent cardiovascular disease, especially coronary artery disease,
CC   cerebrovascular disease, coronary restenosis or peripheral vascular
CC   disease. They may also be used in the treatment of diseases associated
CC   with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC   disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC   The invention specifically excludes proteins with the exact amino acid
CC   sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
CC   acid with the exact sequence as GenBank Accession No: AJ012376.1. The
CC   present sequence represents cDNA encoding a mutant human ABC1 cholesterol
CC   transporter associated with an altered cholesterol level and therefore an
CC   altered risk of cardiovascular disease.
CC   Note: The present sequence is not shown in the specification, but is
CC   derived from the native human ABC1 cDNA shown on pages 157-160.
XX
SQ   Sequence 7864 BP; 2014 A; 1859 C; 2011 G; 1979 T; 1 other;


  Query Match             37.7%;  Score 60;  DB 21;  Length 7864;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-20;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      81 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 140


RESULT  21
AAS06120
ID   AAS06120 standard; cDNA; 9741 BP.
XX
AC   AAS06120;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human ABC1 DNA sequence #1.
XX
KW   Human; ABC1 gene; atherosclerosis; reverse transport; cholesterol;
KW   cardiovascular; neurological; Tangier disease;, LCAT deficiency;
KW   lecithin-cholesterol acetyltransferase; malaria; diabetes; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             185..6967
FT                   /*tag=  a
FT                   /product= "Human ABC1 protein"
XX
PN   WO200130848-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-EP10886.
XX
PR   26-OCT-1999;   99EP-0402668.
PR   01-MAR-2000; 2000US-0186260.
XX
PA   (AVET ) AVENTIS PHARMA SA.
XX
PI   Denefle P,  Rosier-Montus M,  Arnould-Reguigne I,  Prades C,  Naudin L;
PI   Lemoine C,  Duverger N,  Jaye M,  Searfoss GH,  Remaley A,  Brewer HB;
PI   Dean M;
XX
DR   WPI; 2001-316327/33.
DR   P-PSDB; AAU02176.
XX
PT   New human ABC1 nucleic acids and polypeptides for treating
PT   atherosclerosis, malaria and diabetes -
XX
PS   Claim 1; Page 204-208; 368pp; English.
XX
CC   The sequence represents the coding sequence #1 of human ABC1. The
CC   nucleic acid sequence, primers and probes derived from the ABC1 sequence,
CC   and polypeptides and vectors are useful for the prevention of
CC   atherosclerosis, in a subject affected by a dysfunction in the reverse
CC   transport of cholesterol. The polypeptide encoded by the ABC1 gene is
CC   useful for screening for an active ingredient for the prevention or
CC   treatment of a disease resulting from dysfunction in the reverse
CC   transport of cholesterol. The nucleic acids and polypeptides are also
CC   useful for treating and preventing cardiovascular and neurological
CC   pathologies, and other diseases e.g. Tangier disease, lecithin-
CC   cholesterol (LCAT) deficiency, malaria and diabetes.
XX
SQ   Sequence 9741 BP; 2650 A; 2180 C; 2290 G; 2620 T; 1 other;


  Query Match             37.7%;  Score 60;  DB 22;  Length 9741;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-20;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     191 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 250


RESULT  22
AAS06121
ID   AAS06121 standard; cDNA; 9854 BP.
XX
AC   AAS06121;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human ABC1 DNA sequence #2.
XX
KW   Human; ABC1 gene; atherosclerosis; reverse transport; cholesterol;
KW   cardiovascular; neurological; Tangier disease;, LCAT deficiency;
KW   lecithin-cholesterol acetyltransferase; malaria; diabetes; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             298..7078
FT                   /*tag=  a
FT                   /product= "Human ABC1 protein"
XX
PN   WO200130848-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-EP10886.
XX
PR   26-OCT-1999;   99EP-0402668.
PR   01-MAR-2000; 2000US-0186260.
XX
PA   (AVET ) AVENTIS PHARMA SA.
XX
PI   Denefle P,  Rosier-Montus M,  Arnould-Reguigne I,  Prades C,  Naudin L;
PI   Lemoine C,  Duverger N,  Jaye M,  Searfoss GH,  Remaley A,  Brewer HB;
PI   Dean M;
```

```
XX
DR   WPI; 2001-316327/33.
DR   P-PSDB; AAU02176.
XX
PT   New human ABC1 nucleic acids and polypeptides for treating
PT   atherosclerosis, malaria and diabetes -
XX
PS   Claim 1; Page 209-213; 368pp; English.
XX
CC   The sequence represents the coding sequence #2 of human ABC1. The
CC   nucleic acid sequence, primers and probes derived from the ABC1 sequence,
CC   and polypeptides and vectors are useful for the prevention of
CC   atherosclerosis, in a subject affected by a dysfunction in the reverse
CC   transport of cholesterol. The polypeptide encoded by the ABC1 gene is
CC   useful for screening for an active ingredient for the prevention or
CC   treatment of a disease resulting from dysfunction in the reverse
CC   transport of cholesterol. The nucleic acids and polypeptides are also
CC   useful for treating and preventing cardiovascular and neurological
CC   pathologies, and other diseases e.g. Tangier disease, lecithin-
CC   cholesterol (LCAT) deficiency, malaria and diabetes.
XX
SQ   Sequence 9854 BP; 2665 A; 2219 C; 2334 G; 2635 T; 1 other;


  Query Match             37.7%;  Score 60;  DB 22;  Length 9854;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-20;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    304 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 363


RESULT  23
AAC69132
ID   AAC69132 standard; DNA; 10545 BP.
XX
AC   AAC69132;
XX
DT   29-JAN-2001  (first entry)
XX
DE   Human ABC1 gene exon 1 (promoter).
XX
KW   Human ABC1 cholesterol transporter; chromosome 9q31; promoter;
KW   ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW   Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW   cardiovascular disease; coronary artery disease; coronary restenosis;
KW   cerebrovascular disease; peripheral vascular disease;
KW   Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW   X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW   prognosis; prophylaxis; drug screening; transgenic animal; ss.
XX
OS   Homo sapiens.
XX
PN   WO200055318-A2.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-IB00532.
XX
PR   15-MAR-1999;   99US-0124702.
PR   08-JUN-1999;   99US-0138048.
PR   17-JUN-1999;   99US-0139600.
PR   01-SEP-1999;   99US-0151977.
XX
PA   (UYBR-) UNIV BRITISH COLUMBIA.
PA   (XENO-) XENON BIORESEARCH INC.
XX
PI   Hayden MR,  Wilson AR,  Pimstone SN;
XX
DR   WPI; 2000-587528/55.
XX
PT   New ABC1 polypeptide is useful for treating diseases associated with
PT   ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
PT   disease and cancer -
XX
PS   Claim 50; Fig 12; 229pp; English.
XX
CC   The invention relates to the human ABC1 cholesterol transporter protein
CC   (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC   a member of the ATP-binding cassette (ABC transporter) superfamily of
CC   proteins, and plays a crucial role in cholesterol transport, particularly
CC   intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC   involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC   located on chromosome 9q31, and mutations in this gene are associated
CC   with two genetic HDL (high density lipoprotein) deficiency disorders,
CC   Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC   are distinguishable in that TD is an autosomal recessive disorder, while
CC   FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC   cholesterol") in the blood correlate with a high risk of cardiovascular
CC   disease, particularly coronary artery disease, but also cerebrovascular
CC   disease, coronary restenosis, and peripheral vascular disease.
CC   Conversely, a high level of HDL has protective effects against
CC   cardiovascular disease. The invention provides genetic constructs and
CC   transgenic cells and non-human animals comprising human ABC1 nucleic
CC   acids, and methods of gene therapy for the treatment or prevention of
CC   cardiovascular disease comprising the administration of an expression
CC   vector encoding ABC1 or an active fragment thereof. The invention also
CC   encompasses compounds which mimic ABC1 activity, compounds which
CC   stimulate ABC1 expression and methods of screening for such compounds.
CC   It further relates to methods for determining whether a patient has an
CC   increased risk for cardiovascular disease due to polymorphisms in the
CC   ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
CC   or prevent cardiovascular disease, especially coronary artery disease,
CC   cerebrovascular disease, coronary restenosis or peripheral vascular
CC   disease. They may also be used in the treatment of diseases associated
CC   with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC   disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC   The invention specifically excludes proteins with the exact amino acid
CC   sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
CC   acid with the exact sequence as GenBank Accession No: AJ012376.1. The
CC   present sequence represents the human ABC1 gene promoter region (exon 1).
XX
SQ   Sequence 10545 BP; 2647 A; 2225 C; 2411 G; 3256 T; 6 other;


  Query Match             37.7%;  Score 60;  DB 21;  Length 10545;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-20;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8240 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 8299


RESULT  24
AAF92831
ID   AAF92831 standard; DNA; 183999 BP.
XX
AC   AAF92831;
XX
DT   17-MAY-2001  (first entry)
XX
DE   Human ABC1 genomic DNA.
XX
KW   High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ds.
XX
OS   Homo sapiens.
XX
PN   WO200115676-A2.
XX
PD   08-MAR-2001.
XX
PF   01-SEP-2000; 2000WO-IB01492.
XX
PR   01-SEP-1999;   99US-0151977.
```

PR 15-MAR-2000; 2000US-0526193.
PR 23-JUN-2000; 2000US-0213958.
XX
PA (UYBR-) UNIV BRITISH COLUMBIA.
PA (XENO-) XENON GENETICS INC.
XX
PI Hayden MR, Brooks-Wilson AR, Pimstone SN, Clee SM;
XX
DR WPI; 2001-244356/25.
XX
PT Treating a lower than normal high density lipoprotein-cholesterol
PT (HDL-C) level, a higher than normal triglyceride level, or a
PT cardiovascular disease, by administering a compound that modulates LXR-
PT or RXR-mediated transcriptional activity -
XX
PS Claim 8; Fig 1; 317pp; English.
XX
CC The present invention relates to a method for treating a patient
CC diagnosed as having a lower than normal high density
CC lipoprotein-cholesterol (HDL-C) level, a higher than normal
CC triglyceride level, or a cardiovascular disease, involving
CC administering a compound that modulates LXR- or RXR-mediated
CC transcriptional activity or ABC1 expression or activity.
CC The LXR gene product may be used in an assay to identify
CC compounds useful for the treatment of a disease or condition selected a
CC lower than normal HDL cholesterol level, a higher than normal
CC triglyceride level, and a cardiovascular disease.
XX
SQ Sequence 183999 BP; 49549 A; 37944 C; 41170 G; 54950 T; 386 other;

Query Match 37.7%; Score 60; DB 22; Length 183999;
Best Local Similarity 100.0%; Pred. No. 1.7e-20;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 53328 tgttggcctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 53387

RESULT 25
AAH07432
ID AAH07432 standard; cDNA; 736 BP.
XX
AC AAH07432;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:4267.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 1; SEQ ID 4267; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 736 BP; 163 A; 199 C; 199 G; 170 T; 5 other;

Query Match 32.1%; Score 51; DB 22; Length 736;
Best Local Similarity 100.0%; Pred. No. 5.3e-16;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 109 cagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
|||||||||||||||||||||||||||||||||||||||||||||||||||
Db 329 cagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 379

RESULT 26
AAH18606
ID AAH18606 standard; cDNA; 1556 BP.
XX
AC AAH18606;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:18808.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX

DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 18808; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 1556 BP; 380 A; 363 C; 399 G; 414 T; 0 other;

```
  Query Match             32.1%;  Score 51;  DB 22;  Length 1556;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-16;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     109 cagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     329 cagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 379
```

RESULT 27
AAF93084
ID AAF93084 standard; DNA; 37 BP.
XX
AC AAF93084;
XX
DT 17-MAY-2001 (first entry)
XX
DE ABC1 polymorphism RFLP oligionucleotide #45.
XX
KW High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ds.
XX
OS Homo sapiens.
XX
PN WO200115676-A2.
XX
PD 08-MAR-2001.
XX
PF 01-SEP-2000; 2000WO-IB01492.
XX
PR 01-SEP-1999; 99US-0151977.
PR 15-MAR-2000; 2000US-0526193.
PR 23-JUN-2000; 2000US-0213958.
XX
PA (UYBR-) UNIV BRITISH COLUMBIA.
PA (XENO-) XENON GENETICS INC.
XX
PI Hayden MR, Brooks-Wilson AR, Pimstone SN, Clee SM;
XX

DR WPI; 2001-244356/25.
XX
PT Treating a lower than normal high density lipoprotein-cholesterol
PT (HDL-C) level, a higher than normal triglyceride level, or a
PT cardiovascular disease, by administering a compound that modulates LXR-
PT or RXR-mediated transcriptional activity -
XX
PS Disclosure; Fig 17; 317pp; English.
XX
CC The present invention relates to a method for treating a patient
CC diagnosed as having a lower than normal high density
CC lipoprotein-cholesterol (HDL-C) level, a higher than normal
CC triglyceride level, or a cardiovascular disease, involving
CC administering a compound that modulates LXR- or RXR-mediated
CC transcriptional activity or ABC1 expression or activity.
CC The LXR gene product may be used in an assay to identify
CC compounds useful for the treatment of a disease or condition selected a
CC lower than normal HDL cholesterol level, a higher than normal
CC triglyceride level, and a cardiovascular disease.
XX
SQ Sequence 37 BP; 4 A; 17 C; 11 G; 5 T; 0 other;

```
  Query Match             22.0%;  Score 35;  DB 22;  Length 37;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-08;
  Matches   35;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      39 gccgctgccttccagggctcccgagccacacgctg 73
           |||||||||||||||||||||||||||||||||||
Db       1 gccgctgccttccagggctcccgagccacacgctg 35
```

RESULT 28
AAC69306
ID AAC69306 standard; DNA; 21 BP.
XX
AC AAC69306;
XX
DT 29-JAN-2001 (first entry)
XX
DE Human ABC1 gene promoter polymorphic site, SEQ ID NO:205.
XX
KW Human ABC1 cholesterol transporter; chromosome 9q31;
KW ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW cardiovascular disease; coronary artery disease; coronary restenosis;
KW cerebrovascular disease; peripheral vascular disease;
KW Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW prognosis; prophylaxis; drug screening; transgenic animal; ds.
XX
OS Homo sapiens.
XX
PN WO200055318-A2.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-IB00532.
XX
PR 15-MAR-1999; 99US-0124702.
PR 08-JUN-1999; 99US-0138048.
PR 17-JUN-1999; 99US-0139600.
PR 01-SEP-1999; 99US-0151977.
XX
PA (UYBR-) UNIV BRITISH COLUMBIA.
PA (XENO-) XENON BIORESEARCH INC.
XX
PI Hayden MR, Wilson AR, Pimstone SN;
XX
DR WPI; 2000-587528/55.
XX
PT New ABC1 polypeptide is useful for treating diseases associated with
PT ABC1 biological activity, e.g. Alzheimer's disease, Huntington's

```
PT   disease and cancer -
XX
PS   Examples; Fig 11; 229pp; English.
XX
CC   The invention relates to the human ABC1 cholesterol transporter protein
CC   (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC   a member of the ATP-binding cassette (ABC transporter) superfamily of
CC   proteins, and plays a crucial role in cholesterol transport, particularly
CC   intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC   involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC   located on chromosome 9q31, and mutations in this gene are associated
CC   with two genetic HDL (high density lipoprotein) deficiency disorders,
CC   Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC   are distinguishable in that TD is an autosomal recessive disorder, while
CC   FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC   cholesterol") in the blood correlate with a high risk of cardiovascular
CC   disease, particularly coronary artery disease, but also cerebrovascular
CC   disease, coronary restenosis, and peripheral vascular disease.
CC   Conversely, a high level of HDL has protective effects against
CC   cardiovascular disease. The invention provides genetic constructs and
CC   transgenic cells and non-human animals comprising human ABC1 nucleic
CC   acids, and methods of gene therapy for the treatment or prevention of
CC   cardiovascular disease comprising the administration of an expression
CC   vector encoding ABC1 or an active fragment thereof. The invention also
CC   encompasses compounds which mimic ABC1 activity, compounds which
CC   stimulate ABC1 expression and methods of screening for such compounds.
CC   It further relates to methods for determining whether a patient has an
CC   increased risk for cardiovascular disease due to polymorphisms in the
CC   ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
CC   or prevent cardiovascular disease, especially coronary artery disease,
CC   cerebrovascular disease, coronary restenosis or peripheral vascular
CC   disease. They may also be used in the treatment of diseases associated
CC   with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC   disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC   The invention specifically excludes proteins with the exact amino acid
CC   sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
CC   acid with the exact sequence as GenBank Accession No: AJ012376.1.
CC   The present sequence represents a polymorphic site of the human ABC1
CC   gene.
XX
SQ   Sequence 21 BP; 2 A; 6 C; 9 G; 4 T; 0 other;


  Query Match             13.2%;  Score 21;  DB 21;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 0.56;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      66 acacgctgggcgtgctggctg 86
           |||||||||||||||||||||
Db       1 acacgctgggcgtgctggctg 21


RESULT  29
AAC69308
ID   AAC69308 standard; DNA; 21 BP.
XX
AC   AAC69308;
XX
DT   29-JAN-2001  (first entry)
XX
DE   Human ABC1 gene promoter polymorphic site, SEQ ID NO:207.
XX
KW   Human ABC1 cholesterol transporter; chromosome 9q31;
KW   ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW   Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW   cardiovascular disease; coronary artery disease; coronary restenosis;
KW   cerebrovascular disease; peripheral vascular disease;
KW   Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW   X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW   prognosis; prophylaxis; drug screening; transgenic animal; ds.
XX
OS   Homo sapiens.
XX
PN   WO200055318-A2.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-IB00532.
XX
PR   15-MAR-1999;   99US-0124702.
PR   08-JUN-1999;   99US-0138048.
PR   17-JUN-1999;   99US-0139600.
PR   01-SEP-1999;   99US-0151977.
XX
PA   (UYBR-) UNIV BRITISH COLUMBIA.
PA   (XENO-) XENON BIORESEARCH INC.
XX
PI   Hayden MR,  Wilson AR,  Pimstone SN;
XX
DR   WPI; 2000-587528/55.
XX
PT   New ABC1 polypeptide is useful for treating diseases associated with
PT   ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
PT   disease and cancer -
XX
PS   Examples; Fig 11; 229pp; English.
XX
CC   The invention relates to the human ABC1 cholesterol transporter protein
CC   (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC   a member of the ATP-binding cassette (ABC transporter) superfamily of
CC   proteins, and plays a crucial role in cholesterol transport, particularly
CC   intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC   involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC   located on chromosome 9q31, and mutations in this gene are associated
CC   with two genetic HDL (high density lipoprotein) deficiency disorders,
CC   Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC   are distinguishable in that TD is an autosomal recessive disorder, while
CC   FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC   cholesterol") in the blood correlate with a high risk of cardiovascular
CC   disease, particularly coronary artery disease, but also cerebrovascular
CC   disease, coronary restenosis, and peripheral vascular disease.
CC   Conversely, a high level of HDL has protective effects against
CC   cardiovascular disease. The invention provides genetic constructs and
CC   transgenic cells and non-human animals comprising human ABC1 nucleic
CC   acids, and methods of gene therapy for the treatment or prevention of
CC   cardiovascular disease comprising the administration of an expression
CC   vector encoding ABC1 or an active fragment thereof. The invention also
CC   encompasses compounds which mimic ABC1 activity, compounds which
CC   stimulate ABC1 expression and methods of screening for such compounds.
CC   It further relates to methods for determining whether a patient has an
CC   increased risk for cardiovascular disease due to polymorphisms in the
CC   ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
CC   or prevent cardiovascular disease, especially coronary artery disease,
CC   cerebrovascular disease, coronary restenosis or peripheral vascular
CC   disease. They may also be used in the treatment of diseases associated
CC   with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC   disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC   The invention specifically excludes proteins with the exact amino acid
CC   sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
CC   acid with the exact sequence as GenBank Accession No: AJ012376.1.
CC   The present sequence represents a polymorphic site of the human ABC1
CC   gene.
XX
SQ   Sequence 21 BP; 3 A; 9 C; 7 G; 2 T; 0 other;


  Query Match             13.2%;  Score 21;  DB 21;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 0.56;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 gaccagccacgggcgtccctg 26
           |||||||||||||||||||||
Db       1 gaccagccacgggcgtccctg 21


RESULT  30
```

```
AAF92946
ID   AAF92946 standard; DNA; 21 BP.
XX
AC   AAF92946;
XX
DT   17-MAY-2001  (first entry)
XX
DE   Polymorphic sequence for ABC1 polymorphic site #17.
XX
KW   High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ds.
XX
OS   Homo sapiens.
XX
PN   WO200115676-A2.
XX
PD   08-MAR-2001.
XX
PF   01-SEP-2000; 2000WO-IB01492.
XX
PR   01-SEP-1999;   99US-0151977.
PR   15-MAR-2000; 2000US-0526193.
PR   23-JUN-2000; 2000US-0213958.
XX
PA   (UYBR-) UNIV BRITISH COLUMBIA.
PA   (XENO-) XENON GENETICS INC.
XX
PI   Hayden MR,  Brooks-Wilson AR,  Pimstone SN,  Clee SM;
XX
DR   WPI; 2001-244356/25.
XX
PT   Treating a lower than normal high density lipoprotein-cholesterol
PT   (HDL-C) level, a higher than normal triglyceride level, or a
PT   cardiovascular disease, by administering a compound that modulates LXR-
PT   or RXR-mediated transcriptional activity -
XX
PS   Disclosure; Fig 4; 317pp; English.
XX
CC   The present invention relates to a method for treating a patient
CC   diagnosed as having a lower than normal high density
CC   lipoprotein-cholesterol (HDL-C) level, a higher than normal
CC   triglyceride level, or a cardiovascular disease, involving
CC   administering a compound that modulates LXR- or RXR-mediated
CC   transcriptional activity or ABC1 expression or activity.
CC   The LXR gene product may be used in an assay to identify
CC   compounds useful for the treatment of a disease or condition selected a
CC   lower than normal HDL cholesterol level, a higher than normal
CC   triglyceride level, and a cardiovascular disease.
XX
SQ   Sequence 21 BP; 3 A; 10 C; 6 G; 2 T; 0 other;


  Query Match             13.2%;  Score 21;  DB 22;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 0.56;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      7 accagccacgggcgtccctgc 27
          |||||||||||||||||||||
Db      1 accagccacgggcgtccctgc 21


RESULT  31
AAF92948
ID   AAF92948 standard; DNA; 21 BP.
XX
AC   AAF92948;
XX
DT   17-MAY-2001  (first entry)
XX
DE   Polymorphic sequence for ABC1 polymorphic site #18.
XX
KW   High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ds.
XX
OS   Homo sapiens.
XX
PN   WO200115676-A2.
XX
PD   08-MAR-2001.
XX
PF   01-SEP-2000; 2000WO-IB01492.
XX
PR   01-SEP-1999;   99US-0151977.
PR   15-MAR-2000; 2000US-0526193.
PR   23-JUN-2000; 2000US-0213958.
XX
PA   (UYBR-) UNIV BRITISH COLUMBIA.
PA   (XENO-) XENON GENETICS INC.
XX
PI   Hayden MR,  Brooks-Wilson AR,  Pimstone SN,  Clee SM;
XX
DR   WPI; 2001-244356/25.
XX
PT   Treating a lower than normal high density lipoprotein-cholesterol
PT   (HDL-C) level, a higher than normal triglyceride level, or a
PT   cardiovascular disease, by administering a compound that modulates LXR-
PT   or RXR-mediated transcriptional activity -
XX
PS   Disclosure; Fig 4; 317pp; English.
XX
CC   The present invention relates to a method for treating a patient
CC   diagnosed as having a lower than normal high density
CC   lipoprotein-cholesterol (HDL-C) level, a higher than normal
CC   triglyceride level, or a cardiovascular disease, involving
CC   administering a compound that modulates LXR- or RXR-mediated
CC   transcriptional activity or ABC1 expression or activity.
CC   The LXR gene product may be used in an assay to identify
CC   compounds useful for the treatment of a disease or condition selected a
CC   lower than normal HDL cholesterol level, a higher than normal
CC   triglyceride level, and a cardiovascular disease.
XX
SQ   Sequence 21 BP; 2 A; 6 C; 9 G; 4 T; 0 other;


  Query Match             13.2%;  Score 21;  DB 22;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 0.56;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     66 acacgctgggcgtgctggctg 86
          |||||||||||||||||||||
Db      1 acacgctgggcgtgctggctg 21


RESULT  32
AAA68004
ID   AAA68004 standard; DNA; 577 BP.
XX
AC   AAA68004;
XX
DT   24-OCT-2000  (first entry)
XX
DE   Pinus radiata PAL nucleotide sequence SEQ ID NO:97.
XX
KW   Plant; lignin; lignin biosynthetic pathway; Eucalyptus grandis;
KW   Pinus radiata; Monterey pine; ds.
XX
OS   Pinus radiata.
XX
PN   WO200022099-A1.
XX
PD   20-APR-2000.
XX
PF   06-OCT-1999;   99WO-NZ00168.
XX
PR   09-OCT-1998;   98US-0169789.
PR   14-JUL-1999;   99US-0143811.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
```

PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Bloksberg LN, Havukkala IJ;
XX
DR WPI; 2000-317962/27.
XX
PT Novel polynucleotide encoding enzymes involved in lignin-biosynthetic
PT pathway useful for producing transgenic plants especially eucalyptus
PT and pine species having altered lignin content, composition and
PT structure -
XX
PS Claim 1; Page 87; 213pp; English.
XX
CC The present invention describes isolated polynucleotides and proteins
CC encoding and representing the enzymes cinnamate 4-hydroxylase (C4H),
CC coumarate 3-hydroxylase (C3H), phenolase (PNL), O-methyl transferase
CC (OMT), cinnamyl alcohol dehydrogenase (CAD), cinnamoyl-CoA reductase
CC (CCR), phenylalanine ammonia-lyase (PAL), 4-coumarate:CoA ligase (4CL),
CC coniferol glucosyl transferase (CGT), coniferin beta-glucosidase (CBG),
CC laccase, peroxidase, ferulate-5-hydroxylase (F5H), alpha-amylase,
CC caffeic acid methyl transferase, caffeoyl CoA methyl transferase,
CC coumerate CoA ligase, cytochrome P450 LXX1A, diphenol oxidase, flavanol
CC glucosyl transferase, flavenoid hydroxylase, and isoflavone reductase,
CC which are involved in the lignin biosynthetic pathway. The
CC polynucleotides can be used for modulating lignin content, lignin
CC composition and the structure of a plant, especially eucalyptus and pine
CC species, and for modifying the activity of an enzyme involved in lignin
CC biosynthetic pathway, and for producing a plant having altered lignin
CC content, composition and structure. They can be used for designing probes
CC and primers useful for detecting similar DNA and RNA sequences in any
CC organism and for PCR amplification. The lignin content can be efficiently
CC modified using the polynucleotides. AAA67908 to AAA68201 and AAB16341 to
CC AAB16449 represent polynucleotide and protein sequences used in the
CC exemplification of the present invention.
XX
SQ Sequence 577 BP; 128 A; 174 C; 165 G; 110 T; 0 other;

Query Match 11.9%; Score 19; DB 21; Length 577;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 70 gctgggcgtgctggctgag 88
||||||||||||||||||
Db 154 gctgggcgtgctggctgag 172

RESULT 33
AAV23916
ID AAV23916 standard; DNA; 624 BP.
XX
AC AAV23916;
XX
DT 31-JUL-1998 (first entry)
XX
DE Plant PAL enzyme DNA sequence.
XX
KW Lignin biosynthetic pathway; eucalyptus; pine; transgenic plant;
KW lignin content; tree processing; cellulose fibre; ss.
XX
OS Pinus radiata.
XX
PN WO9811205-A2.
XX
PD 19-MAR-1998.
XX
PF 10-SEP-1997; 97WO-NZ00112.
XX
PR 11-SEP-1996; 96US-0713000.
XX
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
PA (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI Bloksberg LN, Grierson AW, Havukkala IJ;
XX
DR WPI; 1998-207374/18.
XX
PT Sequences useful for modification of plant lignin content or
PT structure - from Eucalyptus grandis (eucalyptus) and Pinus radiata
PT (pine) are associated with lignin biosynthesis pathway, useful e.g.
PT in paper industry
XX
PS Example 2; Page 32; 82pp; English.
XX
CC This sequence represents a fragment of the PAL enzyme coding sequence. It
CC is an example of a DNA sequence of the invention, which are from
CC Eucalyptus grandis (eucalyptus) and Pinus radiata (pine) associated with
CC the lignin biosynthesis pathway. Constructs containing the DNA sequences
CC can be used to produce transgenic plants or plant cells, especially woody
CC plants e.g. eucalyptus or pine species but also e.g. monocotyledons or
CC dicotyledons; by stably incorporating the constructs into the plant
CC genome. The lignin content or structure, or activity of a specific enzyme
CC in the plant, can therefore be modulated. Reductions in lignin content or
CC changes in composition are useful in tree processing for paper. High
CC lignin content results in energy- and chemical-intensive separation
CC methods in order to obtain the pure cellulose fibre required. Reductions
CC in lignin content may also be useful for forage crops, whilst increases
CC or changes in composition may be desirable to increase the mechanical
CC strength of wood, change its colour or increase its resistance to rot.
CC The sequences are also useful as probes to isolate DNA sequences encoding
CC enzymes involved in the lignin biosynthesis pathway from other plant
CC species.
XX
SQ Sequence 624 BP; 136 A; 188 C; 188 G; 111 T; 1 other;

Query Match 11.9%; Score 19; DB 19; Length 624;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 70 gctgggcgtgctggctgag 88
||||||||||||||||||
Db 261 gctgggcgtgctggctgag 279

RESULT 34
AAZ06895
ID AAZ06895 standard; cDNA; 624 BP.
XX
AC AAZ06895;
XX
DT 09-NOV-1999 (first entry)
XX
DE Pine phenylalanine ammonia-lyase (PAL) partial cDNA 1.
XX
KW Lignin; biosynthesis; forage crop; wood; paper production;
KW transgenic plant; ss.
XX
OS Pinus radiata.
XX
PN US5952486-A.
XX
PD 14-SEP-1999.
XX
PF 21-NOV-1997; 97US-0975316.
XX
PR 21-NOV-1997; 97US-0975316.
PR 11-SEP-1996; 96US-0713000.
XX
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
PA (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI Bloksberg LN, Grierson AW, Havukkala I;
XX
DR WPI; 1999-527029/44.
XX

PT Isolated DNA sequence encoding enzymes from the lignin synthetic
PT pathway useful for generating plants with an altered lignin content
XX
PS Example 2; Columns 25-26; 48pp; English.
XX
CC This sequence represents a phenylalanine ammonia-lyase (PAL)
CC partial cDNA from Pinus radiata. This enzyme is involved in the
CC biosynthesis of lignin, an insoluble polymer which is primarily
CC responsible for the rigidity of plant stems. Lignin serves as a matrix
CC around the polysaccharide components of some plant cell walls. The
CC higher the lignin content, the more rigid the plant. Lignin also plays a
CC role in disease resistance of plants by impeding the penetration and
CC propagation of pathogenic agents. Lignin is formed by polymerisation of
CC at least three different monolignols (para-coumaryl alcohol, coniferyl
CC alcohol and sinapyl alcohol). These three monolignols are synthesised by
CC similar pathways from phenylalanine in a multistep process and are
CC believed to be polymerised into lignin via a free radical mechanism.
CC The lignin content of plants can be altered using DNA sequences encoding
CC these enzymes. Lignin content can be increased by incorporation of
CC additional copies of genes encoding these enzymes into the target plant.
CC This could be beneficial for increasing the mechanical strength of wood.
CC Similarly, a decrease in lignin content can be obtained by transforming
CC the target plant with antisense copies of such genes. This may be
CC beneficial in plants used as forage crops for livestock (lignin is
CC indigestible) and in trees used in paper manufacture.
XX
SQ Sequence 624 BP; 136 A; 188 C; 188 G; 111 T; 1 other;

Query Match 11.9%; Score 19; DB 20; Length 624;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 70 gctgggcgtgctggctgag 88
||||||||||||||||||
Db 261 gctgggcgtgctggctgag 279

RESULT 35
AAA67916
ID AAA67916 standard; DNA; 624 BP.
XX
AC AAA67916;
XX
DT 24-OCT-2000 (first entry)
XX
DE Pinus radiata PAL nucleotide sequence SEQ ID NO:9.
XX
KW Plant; lignin; lignin biosynthetic pathway; Eucalyptus grandis;
KW Pinus radiata; Monterey pine; ds.
XX
OS Pinus radiata.
XX
PN WO200022099-A1.
XX
PD 20-APR-2000.
XX
PF 06-OCT-1999; 99WO-NZ00168.
XX
PR 09-OCT-1998; 98US-0169789.
PR 14-JUL-1999; 99US-0143811.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Bloksberg LN, Havukkala IJ;
XX
DR WPI; 2000-317962/27.
XX
PT Novel polynucleotide encoding enzymes involved in lignin-biosynthetic
PT pathway useful for producing transgenic plants especially eucalyptus
PT and pine species having altered lignin content, composition and
PT structure -
XX
PS Example 2; Page 59-60; 213pp; English.
XX
CC The present invention describes isolated polynucleotides and proteins
CC encoding and representing the enzymes cinnamate 4-hydroxylase (C4H),
CC coumarate 3-hydroxylase (C3H), phenolase (PNL), O-methyl transferase
CC (OMT), cinnamyl alcohol dehydrogenase (CAD), cinnamoyl-CoA reductase
CC (CCR), phenylalanine ammonia-lyase (PAL), 4-coumarate:CoA ligase (4CL),
CC coniferol glucosyl transferase (CGT), coniferin beta-glucosidase (CBG),
CC laccase, peroxidase, ferulate-5-hydroxylase (F5H), alpha-amylase,
CC caffeic acid methyl transferase, caffeoyl CoA methyl transferase,
CC coumerate CoA ligase, cytochrome P450 LXX1A, diphenol oxidase, flavanol
CC glucosyl transferase, flavenoid hydroxylase, and isoflavone reductase,
CC which are involved in the lignin biosynthetic pathway. The
CC polynucleotides can be used for modulating lignin content, lignin
CC composition and the structure of a plant, especially eucalyptus and pine
CC species, and for modifying the activity of an enzyme involved in lignin
CC biosynthetic pathway, and for producing a plant having altered lignin
CC content, composition and structure. They can be used for designing probes
CC and primers useful for detecting similar DNA and RNA sequences in any
CC organism and for PCR amplification. The lignin content can be efficiently
CC modified using the polynucleotides. AAA67908 to AAA68201 and AAB16341 to
CC AAB16449 represent polynucleotide and protein sequences used in the
CC exemplification of the present invention.
XX
SQ Sequence 624 BP; 136 A; 188 C; 188 G; 111 T; 1 other;

Query Match 11.9%; Score 19; DB 21; Length 624;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 70 gctgggcgtgctggctgag 88
||||||||||||||||||
Db 261 gctgggcgtgctggctgag 279

RESULT 36
AAV23865
ID AAV23865 standard; DNA; 684 BP.
XX
AC AAV23865;
XX
DT 31-JUL-1998 (first entry)
XX
DE Plant PAL enzyme DNA sequence.
XX
KW Lignin biosynthetic pathway; eucalyptus; pine; transgenic plant;
KW lignin content; tree processing; cellulose fibre; ss.
XX
OS Pinus radiata.
XX
PN WO9811205-A2.
XX
PD 19-MAR-1998.
XX
PF 10-SEP-1997; 97WO-NZ00112.
XX
PR 11-SEP-1996; 96US-0713000.
XX
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
PA (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI Bloksberg LN, Grierson AW, Havukkala IJ;
XX
DR WPI; 1998-207374/18.
XX
PT Sequences useful for modification of plant lignin content or
PT structure - from Eucalyptus grandis (eucalyptus) and Pinus radiata
PT (pine) are associated with lignin biosynthesis pathway, useful e.g.
PT in paper industry
XX
PS Claim 1; Page 47; 82pp; English.

```
XX
CC   This sequence represents a fragment of the PAL enzyme coding sequence. It
CC   is an example of a DNA sequence of the invention, which are from
CC   Eucalyptus grandis (eucalyptus) and Pinus radiata (pine) associated with
CC   the lignin biosynthesis pathway. Constructs containing the DNA sequences
CC   can be used to produce transgenic plants or plant cells, especially woody
CC   plants e.g. eucalyptus or pine species but also e.g. monocotyledons or
CC   dicotyledons; by stably incorporating the constructs into the plant
CC   genome. The lignin content or structure, or activity of a specific enzyme
CC   in the plant, can therefore be modulated. Reductions in lignin content or
CC   changes in composition are useful in tree processing for paper. High
CC   lignin content results in energy- and chemical-intensive separation
CC   methods in order to obtain the pure cellulose fibre required. Reductions
CC   in lignin content may also be useful for forage crops, whilst increases
CC   or changes in composition may be desirable to increase the mechanical
CC   strength of wood, change its colour or increase its resistance to rot.
CC   The sequences are also useful as probes to isolate DNA sequences encoding
CC   enzymes involved in the lignin biosynthesis pathway from other plant
CC   species.
XX
SQ   Sequence 684 BP; 150 A; 207 C; 200 G; 127 T; 0 other;


  Query Match             11.9%;  Score 19;  DB 19;  Length 684;
  Best Local Similarity   100.0%;  Pred. No. 5.7;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     70 gctgggcgtgctggctgag 88
          ||||||||||||||||||
Db    261 gctgggcgtgctggctgag 279


RESULT  37
AAZ06898
ID   AAZ06898 standard; cDNA; 684 BP.
XX
AC   AAZ06898;
XX
DT   09-NOV-1999  (first entry)
XX
DE   Pine phenylalanine ammonia-lyase (PAL) partial cDNA 4.
XX
KW   Lignin; biosynthesis; forage crop; wood; paper production;
KW   transgenic plant; ss.
XX
OS   Pinus radiata.
XX
PN   US5952486-A.
XX
PD   14-SEP-1999.
XX
PF   21-NOV-1997;   97US-0975316.
XX
PR   21-NOV-1997;   97US-0975316.
PR   11-SEP-1996;   96US-0713000.
XX
PA   (FLET-) FLETCHER CHALLENGE FORESTS LTD.
PA   (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI   Bloksberg LN,  Grierson AW,  Havukkala I;
XX
DR   WPI; 1999-527029/44.
XX
PT   Isolated DNA sequence encoding enzymes from the lignin synthetic
PT   pathway useful for generating plants with an altered lignin content
XX
PS   Example 2; Columns 49-52; 48pp; English.
XX
CC   This sequence represents a phenylalanine ammonia-lyase (PAL)
CC   partial cDNA from Pinus radiata. This enzyme is involved in the
CC   biosynthesis of lignin, an insoluble polymer which is primarily
CC   responsible for the rigidity of plant stems. Lignin serves as a matrix
CC   around the polysaccharide components of some plant cell walls. The
CC   higher the lignin content, the more rigid the plant. Lignin also plays a
CC   role in disease resistance of plants by impeding the penetration and
CC   propagation of pathogenic agents. Lignin is formed by polymerisation of
CC   at least three different monolignols (para-coumaryl alcohol, coniferyl
CC   alcohol and sinapyl alcohol). These three monolignols are synthesised by
CC   similar pathways from phenylalanine in a multistep process and are
CC   believed to be polymerised into lignin via a free radical mechanism.
CC   The lignin content of plants can be altered using DNA sequences encoding
CC   these enzymes. Lignin content can be increased by incorporation of
CC   additional copies of genes encoding these enzymes into the target plant.
CC   This could be beneficial for increasing the mechanical strength of wood.
CC   Similarly, a decrease in lignin content can be obtained by transforming
CC   the target plant with antisense copies of such genes. This may be
CC   beneficial in plants used as forage crops for livestock (lignin is
CC   indigestible) and in trees used in paper manufacture.
XX
SQ   Sequence 684 BP; 150 A; 207 C; 200 G; 127 T; 0 other;


  Query Match             11.9%;  Score 19;  DB 20;  Length 684;
  Best Local Similarity   100.0%;  Pred. No. 5.7;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     70 gctgggcgtgctggctgag 88
          ||||||||||||||||||
Db    261 gctgggcgtgctggctgag 279


RESULT  38
AAA69586
ID   AAA69586 standard; cDNA; 684 BP.
XX
AC   AAA69586;
XX
DT   08-NOV-2000  (first entry)
XX
DE   Pinus radiata phenylalanine ammonia-lyase cDNA SEQ ID NO:60.
XX
KW   Eucalyptus grandis; Pinus radiata; modification; isoprenoid; plant;
KW   metabolism; isoprenoid biosynthetic pathway; terpenoid; steroid;
KW   genome mapping; physical mapping; positional cloning; forestry;
KW   agriculture; medicine; fermentation; plant development; pest resistance;
KW   pinene; myrcene; Monterey pine; ss.
XX
OS   Pinus radiata.
XX
PN   WO200036081-A2.
XX
PD   22-JUN-2000.
XX
PF   16-DEC-1999;   99WO-NZ00219.
XX
PR   17-DEC-1998;   98US-0215504.
PR   29-JUL-1999;   99US-0146441.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
PA   (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI   Havukkala IJ;
XX
DR   WPI; 2000-431575/37.
XX
PT   New plant polynucleotides encoding polypeptides involved in the
PT   production and modification of isoprenoids, useful in forestry and
PT   agriculture for manipulation of isoprenoid metabolism -
XX
PS   Example 4; Page 73; 164pp; English.
XX
CC   The present invention describes plant polynucleotides encoding
CC   polypeptides involved in the production and modification of isoprenoids,
CC   such as terpenoid and steroid compounds. The polynucleotides are used
CC   in genome mapping, in physical mapping and in positional cloning of
CC   genes. The polynucleotides and polypeptides are useful in forestry and
```

CC agriculture for manipulation of isoprenoid metabolism, in medicine for
CC therapeutic effects, including direct application in diseased organisms
CC or indirect application by transgenic organisms and in fermentation and
CC chemical processing industries involving isoprenoids. In plant
CC applications, manipulating isoprenoid pathways or isoprenoid composition
CC may, for example, affect plant development, pest resistance, and the
CC value of extractives (e.g. pinene and myrcene). The ubiquitous and
CC varied roles of isoprenoids make the polynucleotides attractive targets
CC for biotechnical applications in a variety of fields. AAA69527 to
CC AAA69690 and AAB18004 to AAB18143 represent Eucalyptus grandis and Pinus
CC radiata polynucleotides and proteins used in the exemplification of the
CC present invention.
XX
SQ Sequence 684 BP; 150 A; 207 C; 200 G; 127 T; 0 other;

Query Match 11.9%; Score 19; DB 21; Length 684;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 70 gctgggcgtgctggctgag 88
||||||||||||||||||
Db 261 gctgggcgtgctggctgag 279

RESULT 39
AAA67952
ID AAA67952 standard; DNA; 684 BP.
XX
AC AAA67952;
XX
DT 24-OCT-2000 (first entry)
XX
DE Pinus radiata PAL nucleotide sequence SEQ ID NO:45.
XX
KW Plant; lignin; lignin biosynthetic pathway; Eucalyptus grandis;
KW Pinus radiata; Monterey pine; ds.
XX
OS Pinus radiata.
XX
PN WO200022099-A1.
XX
PD 20-APR-2000.
XX
PF 06-OCT-1999; 99WO-NZ00168.
XX
PR 09-OCT-1998; 98US-0169789.
PR 14-JUL-1999; 99US-0143811.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Bloksberg LN, Havukkala IJ;
XX
DR WPI; 2000-317962/27.
XX
PT Novel polynucleotide encoding enzymes involved in lignin-biosynthetic
PT pathway useful for producing transgenic plants especially eucalyptus
PT and pine species having altered lignin content, composition and
PT structure -
XX
PS Example 2; Page 69; 213pp; English.
XX
CC The present invention describes isolated polynucleotides and proteins
CC encoding and representing the enzymes cinnamate 4-hydroxylase (C4H),
CC coumarate 3-hydroxylase (C3H), phenolase (PNL), O-methyl transferase
CC (OMT), cinnamyl alcohol dehydrogenase (CAD), cinnamoyl-CoA reductase
CC (CCR), phenylalanine ammonia-lyase (PAL), 4-coumarate:CoA ligase (4CL),
CC coniferol glucosyl transferase (CGT), coniferin beta-glucosidase (CBG),
CC laccase, peroxidase, ferulate-5-hydroxylase (F5H), alpha-amylase,
CC caffeic acid methyl transferase, caffeoyl CoA methyl transferase,
CC coumerate CoA ligase, cytochrome P450 LXX1A, diphenol oxidase, flavanol
CC glucosyl transferase, flavenoid hydroxylase, and isoflavone reductase,
CC which are involved in the lignin biosynthetic pathway. The
CC polynucleotides can be used for modulating lignin content, lignin
CC composition and the structure of a plant, especially eucalyptus and pine
CC species, and for modifying the activity of an enzyme involved in lignin
CC biosynthetic pathway, and for producing a plant having altered lignin
CC content, composition and structure. They can be used for designing probes
CC and primers useful for detecting similar DNA and RNA sequences in any
CC organism and for PCR amplification. The lignin content can be efficiently
CC modified using the polynucleotides. AAA67908 to AAA68201 and AAB16341 to
CC AAB16449 represent polynucleotide and protein sequences used in the
CC exemplification of the present invention.
XX
SQ Sequence 684 BP; 150 A; 207 C; 200 G; 127 T; 0 other;

Query Match 11.9%; Score 19; DB 21; Length 684;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 70 gctgggcgtgctggctgag 88
||||||||||||||||||
Db 261 gctgggcgtgctggctgag 279

RESULT 40
AAS73156
ID AAS73156 standard; cDNA; 5286 BP.
XX
AC AAS73156;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #8960.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG08969.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 8960; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating

CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA-and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 5286 BP; 1786 A; 1178 C; 1073 G; 1249 T; 0 other;

Query Match 11.9%; Score 19; DB 23; Length 5286;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 141 tttcagaagaagacaaaca 159
||||||||||||||||||||
Db 1961 tttcagaagaagacaaaca 1979

RESULT 41
AAS80591
ID AAS80591 standard; cDNA; 5954 BP.
XX
AC AAS80591;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #16395.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG16404.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 16395; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 5954 BP; 2007 A; 1302 C; 1219 G; 1426 T; 0 other;

Query Match 11.9%; Score 19; DB 23; Length 5954;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 141 tttcagaagaagacaaaca 159
||||||||||||||||||||
Db 2104 tttcagaagaagacaaaca 2122

RESULT 42
AAS83843/c
ID AAS83843 standard; cDNA; 6143 BP.
XX
AC AAS83843;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #19647.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG19656.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 19647; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in

```
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 6143 BP; 1489 A; 1251 C; 1342 G; 2061 T; 0 other;


  Query Match             11.9%;  Score 19;  DB 23;  Length 6143;
  Best Local Similarity   100.0%;  Pred. No. 5.8;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    141 tttcagaagaagacaaaca 159
          ||||||||||||||||||||
Db   4040 TTTCAGAAGAAGACAAACA 4022


RESULT  43
AAC69153/c
ID   AAC69153 standard; DNA; 18 BP.
XX
AC   AAC69153;
XX
DT   29-JAN-2001  (first entry)
XX
DE   Human ABC1 phosphorothioate antisense oligonucleotide AN-7.
XX
KW   Human ABC1 cholesterol transporter; chromosome 9q31;
KW   ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW   Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW   cardiovascular disease; coronary artery disease; coronary restenosis;
KW   cerebrovascular disease; peripheral vascular disease;
KW   Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW   X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW   prognosis; prophylaxis; drug screening; transgenic animal;
KW   phosphorothioate antisense oligonucleotide; ss.
XX
OS   Homo sapiens.
XX
PN   WO200055318-A2.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-IB00532.
XX
PR   15-MAR-1999;   99US-0124702.
PR   08-JUN-1999;   99US-0138048.
PR   17-JUN-1999;   99US-0139600.
PR   01-SEP-1999;   99US-0151977.
XX
PA   (UYBR-) UNIV BRITISH COLUMBIA.
PA   (XENO-) XENON BIORESEARCH INC.
XX
PI   Hayden MR,  Wilson AR,  Pimstone SN;
XX
DR   WPI; 2000-587528/55.
XX
PT   New ABC1 polypeptide is useful for treating diseases associated with
PT   ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
PT   disease and cancer -
XX
PS   Examples; Page 39; 229pp; English.
XX
CC   The invention relates to the human ABC1 cholesterol transporter protein
CC   (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC   a member of the ATP-binding cassette (ABC transporter) superfamily of
CC   proteins, and plays a crucial role in cholesterol transport, particularly
CC   intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC   involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC   located on chromosome 9q31, and mutations in this gene are associated
CC   with two genetic HDL (high density lipoprotein) deficiency disorders,
CC   Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC   are distinguishable in that TD is an autosomal recessive disorder, while
CC   FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC   cholesterol") in the blood correlate with a high risk of cardiovascular
CC   disease, particularly coronary artery disease, but also cerebrovascular
CC   disease, coronary restenosis, and peripheral vascular disease.
CC   Conversely, a high level of HDL has protective effects against
CC   cardiovascular disease. The invention provides genetic constructs and
CC   transgenic cells and non-human animals comprising human ABC1 nucleic
CC   acids, and methods of gene therapy for the treatment or prevention of
CC   cardiovascular disease comprising the administration of an expression
CC   vector encoding ABC1 or an active fragment thereof. The invention also
CC   encompasses compounds which mimic ABC1 activity, compounds which
CC   stimulate ABC1 expression and methods of screening for such compounds.
CC   It further relates to methods for determining whether a patient has an
CC   increased risk for cardiovascular disease due to polymorphisms in the
CC   ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
CC   or prevent cardiovascular disease, especially coronary artery disease,
CC   cerebrovascular disease, coronary restenosis or peripheral vascular
CC   disease. They may also be used in the treatment of diseases associated
CC   with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC   disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC   The invention specifically excludes proteins with the exact amino acid
CC   sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
CC   acid with the exact sequence as GenBank Accession No: AJ012376.1. The
CC   present sequence represents a human ABC-1 cDNA-specific phosphorothioate
CC   antisense oligonucleotide used in the exemplifications of the invention.
XX
SQ   Sequence 18 BP; 3 A; 8 C; 3 G; 4 T; 0 other;


  Query Match             11.3%;  Score 18;  DB 21;  Length 18;
  Best Local Similarity   100.0%;  Pred. No. 18;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     79 gctggctgagggaacatg 96
          ||||||||||||||||||
Db     18 GCTGGCTGAGGGAACATG 1


RESULT  44
ABL08833/c
ID   ABL08833 standard; cDNA; 6420 BP.
XX
AC   ABL08833;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster expressed polynucleotide SEQ ID NO 20981.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ss.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   P-PSDB; ABB64730.
XX
```

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 20981; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 6420 BP; 1886 A; 1829 C; 1603 G; 1102 T; 0 other;

Query Match 11.3%; Score 18; DB 23; Length 6420;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 110 agctgaggttgctgctgt 127
||||||||||||||||||
Db 4817 AGCTGAGGTTGCTGCTGT 4800

RESULT 45
ABL08832/c
ID ABL08832 standard; cDNA; 11580 BP.
XX
AC ABL08832;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 20978.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR P-PSDB; ABB64729.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 20978; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 11580 BP; 3621 A; 2752 C; 2467 G; 2740 T; 0 other;

Query Match 11.3%; Score 18; DB 23; Length 11580;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 110 agctgaggttgctgctgt 127
||||||||||||||||||
Db 6963 AGCTGAGGTTGCTGCTGT 6946

Search completed: September 20, 2002, 06:08:29
Job time: 10393 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 06:15:34 ; Search time 111.56 Seconds
(without alignments)
350.087 Million cell updates/sec

Title: US-09-846-456-5
Perfect score: 159
Sequence: 1 ttaatgaccagccacgggcg.....ctttcagaagaagacaaaca 159

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	19	11.9	624	2	US-08-713-000-9	Sequence 9, Appli
	2	19	11.9	624	2	US-08-975-316-9	Sequence 9, Appli
	3	19	11.9	624	4	US-09-211-710-9	Sequence 9, Appli
	4	19	11.9	684	2	US-08-975-316-45	Sequence 45, Appl
c	5	16	10.1	39	1	US-08-519-777-63	Sequence 63, Appl
c	6	16	10.1	39	1	US-08-742-035-63	Sequence 63, Appl
c	7	16	10.1	39	1	US-08-777-019-63	Sequence 63, Appl
c	8	16	10.1	39	2	US-08-777-143-63	Sequence 63, Appl
c	9	16	10.1	39	3	US-08-775-414-63	Sequence 63, Appl
c	10	16	10.1	39	4	US-08-931-858E-63	Sequence 63, Appl
c	11	16	10.1	39	4	US-08-981-739-63	Sequence 63, Appl
c	12	16	10.1	57	1	US-08-519-777-18	Sequence 18, Appl
c	13	16	10.1	57	1	US-08-742-035-18	Sequence 18, Appl
c	14	16	10.1	57	1	US-08-777-019-18	Sequence 18, Appl
c	15	16	10.1	57	2	US-08-777-143-18	Sequence 18, Appl
c	16	16	10.1	57	3	US-08-775-414-18	Sequence 18, Appl
c	17	16	10.1	57	4	US-08-931-858E-18	Sequence 18, Appl
c	18	16	10.1	57	4	US-08-981-739-18	Sequence 18, Appl
c	19	16	10.1	169	1	US-08-519-777-29	Sequence 29, Appl
c	20	16	10.1	169	1	US-08-742-035-29	Sequence 29, Appl
c	21	16	10.1	169	1	US-08-777-019-29	Sequence 29, Appl
c	22	16	10.1	169	2	US-08-777-143-29	Sequence 29, Appl
c	23	16	10.1	169	3	US-08-775-414-29	Sequence 29, Appl
c	24	16	10.1	169	4	US-08-931-858E-29	Sequence 29, Appl
c	25	16	10.1	169	4	US-08-981-739-29	Sequence 29, Appl
c	26	16	10.1	285	1	US-08-519-777-26	Sequence 26, Appl
c	27	16	10.1	285	1	US-08-742-035-26	Sequence 26, Appl
c	28	16	10.1	285	1	US-08-777-019-26	Sequence 26, Appl
c	29	16	10.1	285	2	US-08-777-143-26	Sequence 26, Appl
c	30	16	10.1	285	3	US-08-775-414-26	Sequence 26, Appl
c	31	16	10.1	285	4	US-08-931-858E-26	Sequence 26, Appl
c	32	16	10.1	285	4	US-08-981-739-26	Sequence 26, Appl
c	33	16	10.1	585	1	US-08-519-777-12	Sequence 12, Appl
c	34	16	10.1	585	1	US-08-742-035-12	Sequence 12, Appl
c	35	16	10.1	585	1	US-08-777-019-12	Sequence 12, Appl
c	36	16	10.1	585	2	US-08-777-143-12	Sequence 12, Appl
c	37	16	10.1	585	3	US-08-775-414-12	Sequence 12, Appl
c	38	16	10.1	585	4	US-08-931-858E-12	Sequence 12, Appl
c	39	16	10.1	585	4	US-08-981-739-12	Sequence 12, Appl
c	40	16	10.1	1023	4	US-08-931-858E-49	Sequence 49, Appl
c	41	16	10.1	1351	4	US-08-697-954-3	Sequence 3, Appli
c	42	16	10.1	1872	1	US-08-428-926-1	Sequence 1, Appli
c	43	16	10.1	1872	1	US-08-435-434-4	Sequence 4, Appli
c	44	16	10.1	1872	1	US-08-435-436-4	Sequence 4, Appli
c	45	16	10.1	1872	1	US-08-428-927-1	Sequence 1, Appli

ALIGNMENTS

```
RESULT   1
US-08-713-000-9
; Sequence 9, Application US/08713000
; Patent No. 5850020
;  GENERAL INFORMATION:
;    APPLICANT:  Bloksberg, Leonard N.
;    APPLICANT:  Havukkala, Ilkka
;    APPLICANT:  Grierson, Alastair
;    TITLE OF INVENTION:  MATERIALS AND METHODS FOR THE
;    TITLE OF INVENTION:  MODIFICATION OF PLANT LIGNIN CONTENT
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Speckman Picard PLLC
;      STREET:  2601 Elliott Avenue, Suite 4185
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/713,000
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sleath, Janet
;      REGISTRATION NUMBER:  37,007
;      REFERENCE/DOCKET NUMBER:  11000.1003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-269-0565
;      TELEFAX:  206-269-0563
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  624 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-713-000-9


  Query Match             11.9%;  Score 19;  DB 2;  Length 624;
  Best Local Similarity   100.0%;  Pred. No. 0.57;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      70 gctgggcgtgctggctgag 88
           ||||||||||||||||||
Db     261 GCTGGGCGTGCTGGCTGAG 279
```

```
RESULT   2
US-08-975-316-9
; Sequence 9, Application US/08975316
; Patent No. 5952486
;  GENERAL INFORMATION:
;    APPLICANT:  BLOKSBERG, Leonard N., HAVUKKALA, Ilkka
;    APPLICANT:  and GRIERSON, Alastair W.
;    TITLE OF INVENTION:  MATERIALS AND METHODS FOR
;    TITLE OF INVENTION:  THE MODIFICATION OF PLANT LIGNIN CONTENT
;    NUMBER OF SEQUENCES:  88
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Law Offices of Ann W. Speckman
;      STREET:  2601 Elliott Avenue, Suite 4185
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/975,316
;      FILING DATE:
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/713,000
;      FILING DATE:  September 11, 1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  SLEATH, Janet
;      REGISTRATION NUMBER:  37,007
;      REFERENCE/DOCKET NUMBER:  11000/1003C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-269-0565
;      TELEFAX:  206-269-0563
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  624 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-975-316-9


  Query Match            11.9%;  Score 19;  DB 2;  Length 624;
  Best Local Similarity  100.0%;  Pred. No. 0.57;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      70 gctgggcgtgctggctgag 88
           ||||||||||||||||||||
Db     261 GCTGGGCGTGCTGGCTGAG 279


RESULT   3
US-09-211-710-9
; Sequence 9, Application US/09211710A
; Patent No. 6204434
; GENERAL INFORMATION:
;  APPLICANT: Bloksberg, Leonard N.
;  APPLICANT: Havukkala, Ilkka
;  APPLICANT: Grierson, Alastair
;  TITLE OF INVENTION: Materials and Methods for the
;  TITLE OF INVENTION: Modification of Plant Lignin Content
;  FILE REFERENCE: 11000.1003c3
;  CURRENT APPLICATION NUMBER: US/09/211,710A
;  CURRENT FILING DATE: 1998-12-14
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 9
;   LENGTH: 624
;   TYPE: DNA
;   ORGANISM: Pinus radiata
US-09-211-710-9


  Query Match            11.9%;  Score 19;  DB 4;  Length 624;
  Best Local Similarity  100.0%;  Pred. No. 0.57;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      70 gctgggcgtgctggctgag 88
           ||||||||||||||||||||
Db     261 gctgggcgtgctggctgag 279


RESULT   4
US-08-975-316-45
; Sequence 45, Application US/08975316
; Patent No. 5952486
;  GENERAL INFORMATION:
;    APPLICANT:  BLOKSBERG, Leonard N., HAVUKKALA, Ilkka
;    APPLICANT:  and GRIERSON, Alastair W.
;    TITLE OF INVENTION:  MATERIALS AND METHODS FOR
;    TITLE OF INVENTION:  THE MODIFICATION OF PLANT LIGNIN CONTENT
;    NUMBER OF SEQUENCES:  88
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Law Offices of Ann W. Speckman
;      STREET:  2601 Elliott Avenue, Suite 4185
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/975,316
;      FILING DATE:
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/713,000
;      FILING DATE:  September 11, 1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  SLEATH, Janet
;      REGISTRATION NUMBER:  37,007
;      REFERENCE/DOCKET NUMBER:  11000/1003C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-269-0565
;      TELEFAX:  206-269-0563
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  45:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  684 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-975-316-45


  Query Match            11.9%;  Score 19;  DB 2;  Length 684;
  Best Local Similarity  100.0%;  Pred. No. 0.57;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      70 gctgggcgtgctggctgag 88
           ||||||||||||||||||||
Db     261 GCTGGGCGTGCTGGCTGAG 279


RESULT   5
US-08-519-777-63/c
; Sequence 63, Application US/08519777
```

```
; Patent No. 5739307
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROGERS, HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/519,777
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  953095
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  63:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  39 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-519-777-63


  Query Match             10.1%;  Score 16;  DB 1;  Length 39;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      36 GGCCGCTGCCTTCCAG 21


RESULT   6
US-08-742-035-63/c
; Sequence 63, Application US/08742035
; Patent No. 5747655
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROGERS, HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/742,035
;      FILING DATE:  01-NOV-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/519,777
;      FILING DATE:  28-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  953095
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  63:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  39 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-742-035-63


  Query Match             10.1%;  Score 16;  DB 1;  Length 39;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      36 GGCCGCTGCCTTCCAG 21


RESULT   7
US-08-777-019-63/c
; Sequence 63, Application US/08777019
; Patent No. 5817622
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROGERS, HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/777,019
;      FILING DATE:  30-DEC-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/519,777
;      FILING DATE:  28-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  953095
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  63:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  39 base pairs
```

```
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-777-019-63


  Query Match             10.1%;  Score 16;  DB 1;  Length 39;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      36 GGCCGCTGCCTTCCAG 21


RESULT   8
US-08-777-143-63/c
; Sequence 63, Application US/08777143
; Patent No. 5843914
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROGERS, HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/777,143
;      FILING DATE:  30-DEC-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/519,777
;      FILING DATE:  28-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  953095
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  63:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  39 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-777-143-63


  Query Match             10.1%;  Score 16;  DB 2;  Length 39;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      36 GGCCGCTGCCTTCCAG 21


RESULT   9
US-08-775-414-63/c
; Sequence 63, Application US/08775414
; Patent No. 6090778
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  90
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/775,414
;      FILING DATE:  31-DEC-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965805
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  63:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  39 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-775-414-63


  Query Match             10.1%;  Score 16;  DB 3;  Length 39;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      36 GGCCGCTGCCTTCCAG 21


RESULT  10
US-08-931-858E-63/c
; Sequence 63, Application US/08931858E
; Patent No. 6222022
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON, EUGENE M
;    APPLICANT:  MILBRANDT, JEFFREY D
;    APPLICANT:  KOTZBAUER, PAUL T
;    APPLICANT:  LAMPE, PATRICIA A
;    APPLICANT:  KLEIN, ROBERT
;    APPLICANT:  DESAUVAGE, FRED
;    TITLE OF INVENTION:  PERSEPHIN AND RELATED GROWTH FACTOR
;    NUMBER OF SEQUENCES:  239
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MO
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
```

```
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/931,858E
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  971486
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  314-727-5188
;      TELEFAX:  314-727-6092
;  INFORMATION FOR SEQ ID NO:  63:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  39 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-931-858E-63


  Query Match             10.1%;  Score 16;  DB 4;  Length 39;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      36 GGCCGCTGCCTTCCAG 21


RESULT  11
US-08-981-739-63/c
; Sequence 63, Application US/08981739
; Patent No. 6232449
;  GENERAL INFORMATION:
;       APPLICANT: JOHNSON JR., EUGENE M.
;                  MILBRANDT, JEFFREY D.
;                  KOTZBAUER, PAUL T.
;                  LAMPE, PATRICIA A.
;       TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTORS
;       NUMBER OF SEQUENCES: 176
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: HOWELL & HAFERKAMP, L.C.
;            STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
;            CITY: ST. LOUIS
;            STATE: MISSOURI
;            COUNTRY: US
;            ZIP: 63105-1817
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/08/981,739
;            FILING DATE: 31-Aug-1998
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: PCT/US97/03461
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: HOLLAND, DONALD R.
;            REGISTRATION NUMBER: 35,197
;            REFERENCE/DOCKET NUMBER: 976163
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (314) 727-5188
;            TELEFAX: (314) 727-6092
;  INFORMATION FOR SEQ ID NO: 63:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 39 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 63:
US-08-981-739-63


  Query Match             10.1%;  Score 16;  DB 4;  Length 39;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      36 GGCCGCTGCCTTCCAG 21


RESULT  12
US-08-519-777-18/c
; Sequence 18, Application US/08519777
; Patent No. 5739307
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROGERS, HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/519,777
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  953095
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  57 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-519-777-18


  Query Match             10.1%;  Score 16;  DB 1;  Length 57;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  13
US-08-742-035-18/c
```

```
; Sequence 18, Application US/08742035
; Patent No. 5747655
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROGERS, HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/742,035
;      FILING DATE:  01-NOV-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/519,777
;      FILING DATE:  28-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  953095
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  57 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-742-035-18


  Query Match             10.1%;  Score 16;  DB 1;  Length 57;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  14
US-08-777-019-18/c
; Sequence 18, Application US/08777019
; Patent No. 5817622
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROGERS, HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/777,019
;      FILING DATE:  30-DEC-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/519,777
;      FILING DATE:  28-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  953095
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  57 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-777-019-18


  Query Match             10.1%;  Score 16;  DB 1;  Length 57;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  15
US-08-777-143-18/c
; Sequence 18, Application US/08777143
; Patent No. 5843914
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROGERS, HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/777,143
;      FILING DATE:  30-DEC-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/519,777
;      FILING DATE:  28-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  953095
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
```

```
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  57 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-777-143-18


  Query Match             10.1%;  Score 16;  DB 2;  Length 57;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  16
US-08-775-414-18/c
; Sequence 18, Application US/08775414
; Patent No. 6090778
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  90
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/775,414
;      FILING DATE:  31-DEC-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965805
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  57 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-775-414-18


  Query Match             10.1%;  Score 16;  DB 3;  Length 57;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  17
US-08-931-858E-18/c
; Sequence 18, Application US/08931858E
; Patent No. 6222022
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON, EUGENE M
;    APPLICANT:  MILBRANDT, JEFFREY D
;    APPLICANT:  KOTZBAUER, PAUL T
;    APPLICANT:  LAMPE, PATRICIA A
;    APPLICANT:  KLEIN, ROBERT
;    APPLICANT:  DESAUVAGE, FRED
;    TITLE OF INVENTION:  PERSEPHIN AND RELATED GROWTH FACTOR
;    NUMBER OF SEQUENCES:  239
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MO
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/931,858E
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  971486
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  314-727-5188
;      TELEFAX:  314-727-6092
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  57 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-931-858E-18


  Query Match             10.1%;  Score 16;  DB 4;  Length 57;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  18
US-08-981-739-18/c
; Sequence 18, Application US/08981739
; Patent No. 6232449
;   GENERAL INFORMATION:
;        APPLICANT: JOHNSON JR., EUGENE M.
;                   MILBRANDT, JEFFREY D.
;                   KOTZBAUER, PAUL T.
;                   LAMPE, PATRICIA A.
;        TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTORS
;        NUMBER OF SEQUENCES: 176
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: HOWELL & HAFERKAMP, L.C.
;             STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
;             CITY: ST. LOUIS
;             STATE: MISSOURI
;             COUNTRY: US
;             ZIP: 63105-1817
```

```
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/981,739
;             FILING DATE: 31-Aug-1998
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: PCT/US97/03461
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: HOLLAND, DONALD R.
;             REGISTRATION NUMBER: 35,197
;             REFERENCE/DOCKET NUMBER: 976163
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (314) 727-5188
;             TELEFAX: (314) 727-6092
;   INFORMATION FOR SEQ ID NO: 18:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 57 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-08-981-739-18


  Query Match             10.1%;  Score 16;  DB 4;  Length 57;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  19
US-08-519-777-29/c
; Sequence 29, Application US/08519777
; Patent No. 5739307
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROGERS, HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/519,777
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  953095
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  169 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-519-777-29


  Query Match             10.1%;  Score 16;  DB 1;  Length 169;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  20
US-08-742-035-29/c
; Sequence 29, Application US/08742035
; Patent No. 5747655
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROGERS, HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/742,035
;      FILING DATE:  01-NOV-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/519,777
;      FILING DATE:  28-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  953095
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  169 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-742-035-29


  Query Match             10.1%;  Score 16;  DB 1;  Length 169;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9
```

```
RESULT  21
US-08-777-019-29/c
; Sequence 29, Application US/08777019
; Patent No. 5817622
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROGERS, HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/777,019
;      FILING DATE:  30-DEC-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/519,777
;      FILING DATE:  28-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  953095
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  169 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-777-019-29


  Query Match             10.1%;  Score 16;  DB 1;  Length 169;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  22
US-08-777-143-29/c
; Sequence 29, Application US/08777143
; Patent No. 5843914
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROGERS, HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/777,143
;      FILING DATE:  30-DEC-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/519,777
;      FILING DATE:  28-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  953095
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  169 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-777-143-29


  Query Match             10.1%;  Score 16;  DB 2;  Length 169;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  23
US-08-775-414-29/c
; Sequence 29, Application US/08775414
; Patent No. 6090778
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  90
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/775,414
;      FILING DATE:  31-DEC-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965805
;    TELECOMMUNICATION INFORMATION:
```

```
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;   INFORMATION FOR SEQ ID NO:  29:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  169 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
US-08-775-414-29


  Query Match             10.1%;  Score 16;  DB 3;  Length 169;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  24
US-08-931-858E-29/c
; Sequence 29, Application US/08931858E
; Patent No. 6222022
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON, EUGENE M
;    APPLICANT:  MILBRANDT, JEFFREY D
;    APPLICANT:  KOTZBAUER, PAUL T
;    APPLICANT:  LAMPE, PATRICIA A
;    APPLICANT:  KLEIN, ROBERT
;    APPLICANT:  DESAUVAGE, FRED
;    TITLE OF INVENTION:  PERSEPHIN AND RELATED GROWTH FACTOR
;    NUMBER OF SEQUENCES:  239
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MO
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/931,858E
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  971486
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  314-727-5188
;      TELEFAX:  314-727-6092
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  169 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-931-858E-29


  Query Match             10.1%;  Score 16;  DB 4;  Length 169;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  25
US-08-981-739-29/c
; Sequence 29, Application US/08981739
; Patent No. 6232449
;  GENERAL INFORMATION:
;       APPLICANT: JOHNSON JR., EUGENE M.
;                  MILBRANDT, JEFFREY D.
;                  KOTZBAUER, PAUL T.
;                  LAMPE, PATRICIA A.
;       TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTORS
;       NUMBER OF SEQUENCES: 176
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: HOWELL & HAFERKAMP, L.C.
;            STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
;            CITY: ST. LOUIS
;            STATE: MISSOURI
;            COUNTRY: US
;            ZIP: 63105-1817
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/08/981,739
;            FILING DATE: 31-Aug-1998
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: PCT/US97/03461
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: HOLLAND, DONALD R.
;            REGISTRATION NUMBER: 35,197
;            REFERENCE/DOCKET NUMBER: 976163
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (314) 727-5188
;            TELEFAX: (314) 727-6092
;   INFORMATION FOR SEQ ID NO: 29:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 169 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 29:
US-08-981-739-29


  Query Match             10.1%;  Score 16;  DB 4;  Length 169;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  26
US-08-519-777-26/c
; Sequence 26, Application US/08519777
; Patent No. 5739307
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROGERS, HOWELL & HAFERKAMP, L.C.
```

```
;       STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;       CITY:  ST. LOUIS
;       STATE:  MISSOURI
;       COUNTRY:  US
;       ZIP:  63105-1817
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/519,777
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  HOLLAND, DONALD R.
;       REGISTRATION NUMBER:  35,197
;       REFERENCE/DOCKET NUMBER:  953095
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (314) 727-5188
;       TELEFAX:  (314) 727-6092
;   INFORMATION FOR SEQ ID NO:  26:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  285 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
US-08-519-777-26


  Query Match             10.1%;  Score 16;  DB 1;  Length 285;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  27
US-08-742-035-26/c
; Sequence 26, Application US/08742035
; Patent No. 5747655
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROGERS, HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/742,035
;      FILING DATE:  01-NOV-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/519,777
;      FILING DATE:  28-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  953095
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  285 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-742-035-26


  Query Match             10.1%;  Score 16;  DB 1;  Length 285;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  28
US-08-777-019-26/c
; Sequence 26, Application US/08777019
; Patent No. 5817622
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROGERS, HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/777,019
;      FILING DATE:  30-DEC-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/519,777
;      FILING DATE:  28-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  953095
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  285 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-777-019-26


  Query Match             10.1%;  Score 16;  DB 1;  Length 285;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  29
US-08-777-143-26/c
; Sequence 26, Application US/08777143
; Patent No. 5843914
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROGERS, HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/777,143
;      FILING DATE:  30-DEC-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/519,777
;      FILING DATE:  28-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  953095
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  285 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-777-143-26


  Query Match             10.1%;  Score 16;  DB 2;  Length 285;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  30
US-08-775-414-26/c
; Sequence 26, Application US/08775414
; Patent No. 6090778
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  90
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/775,414
;      FILING DATE:  31-DEC-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965805
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  285 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-775-414-26


  Query Match             10.1%;  Score 16;  DB 3;  Length 285;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  31
US-08-931-858E-26/c
; Sequence 26, Application US/08931858E
; Patent No. 6222022
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON, EUGENE M
;    APPLICANT:  MILBRANDT, JEFFREY D
;    APPLICANT:  KOTZBAUER, PAUL T
;    APPLICANT:  LAMPE, PATRICIA A
;    APPLICANT:  KLEIN, ROBERT
;    APPLICANT:  DESAUVAGE, FRED
;    TITLE OF INVENTION:  PERSEPHIN AND RELATED GROWTH FACTOR
;    NUMBER OF SEQUENCES:  239
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MO
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/931,858E
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
```

```
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  971486
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  314-727-5188
;      TELEFAX:  314-727-6092
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  285 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-931-858E-26


  Query Match           10.1%;  Score 16;  DB 4;  Length 285;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  32
US-08-981-739-26/c
; Sequence 26, Application US/08981739
; Patent No. 6232449
;  GENERAL INFORMATION:
;       APPLICANT: JOHNSON JR., EUGENE M.
;                  MILBRANDT, JEFFREY D.
;                  KOTZBAUER, PAUL T.
;                  LAMPE, PATRICIA A.
;       TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTORS
;       NUMBER OF SEQUENCES: 176
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: HOWELL & HAFERKAMP, L.C.
;            STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
;            CITY: ST. LOUIS
;            STATE: MISSOURI
;            COUNTRY: US
;            ZIP: 63105-1817
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/08/981,739
;            FILING DATE: 31-Aug-1998
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: PCT/US97/03461
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: HOLLAND, DONALD R.
;            REGISTRATION NUMBER: 35,197
;            REFERENCE/DOCKET NUMBER: 976163
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (314) 727-5188
;            TELEFAX: (314) 727-6092
;   INFORMATION FOR SEQ ID NO: 26:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 285 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-08-981-739-26


  Query Match           10.1%;  Score 16;  DB 4;  Length 285;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  33
US-08-519-777-12/c
; Sequence 12, Application US/08519777
; Patent No. 5739307
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROGERS, HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/519,777
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  953095
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  585 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-519-777-12


  Query Match           10.1%;  Score 16;  DB 1;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  34
US-08-742-035-12/c
; Sequence 12, Application US/08742035
; Patent No. 5747655
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
```

```
;      ADDRESSEE:  ROGERS, HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/742,035
;      FILING DATE:  01-NOV-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/519,777
;      FILING DATE:  28-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  953095
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  585 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-742-035-12


  Query Match             10.1%;  Score 16;  DB 1;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  35
US-08-777-019-12/c
; Sequence 12, Application US/08777019
; Patent No. 5817622
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROGERS, HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/777,019
;      FILING DATE:  30-DEC-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/519,777
;      FILING DATE:  28-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  953095
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  585 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-777-019-12


  Query Match             10.1%;  Score 16;  DB 1;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  36
US-08-777-143-12/c
; Sequence 12, Application US/08777143
; Patent No. 5843914
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROGERS, HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/777,143
;      FILING DATE:  30-DEC-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/519,777
;      FILING DATE:  28-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  953095
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  585 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-777-143-12
```

```
  Query Match             10.1%;  Score 16;  DB 2;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  37
US-08-775-414-12/c
; Sequence 12, Application US/08775414
; Patent No. 6090778
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  90
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/775,414
;      FILING DATE:  31-DEC-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965805
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  585 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-775-414-12


  Query Match             10.1%;  Score 16;  DB 3;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  38
US-08-931-858E-12/c
; Sequence 12, Application US/08931858E
; Patent No. 6222022
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON, EUGENE M
;    APPLICANT:  MILBRANDT, JEFFREY D
;    APPLICANT:  KOTZBAUER, PAUL T
;    APPLICANT:  LAMPE, PATRICIA A
;    APPLICANT:  KLEIN, ROBERT
;    APPLICANT:  DESAUVAGE, FRED
;    TITLE OF INVENTION:  PERSEPHIN AND RELATED GROWTH FACTOR
;    NUMBER OF SEQUENCES:  239
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MO
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/931,858E
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  971486
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  314-727-5188
;      TELEFAX:  314-727-6092
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  585 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-931-858E-12


  Query Match             10.1%;  Score 16;  DB 4;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  39
US-08-981-739-12/c
; Sequence 12, Application US/08981739
; Patent No. 6232449
;   GENERAL INFORMATION:
;        APPLICANT: JOHNSON JR., EUGENE M.
;                   MILBRANDT, JEFFREY D.
;                   KOTZBAUER, PAUL T.
;                   LAMPE, PATRICIA A.
;        TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTORS
;        NUMBER OF SEQUENCES: 176
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: HOWELL & HAFERKAMP, L.C.
;             STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
;             CITY: ST. LOUIS
;             STATE: MISSOURI
;             COUNTRY: US
;             ZIP: 63105-1817
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/981,739
;             FILING DATE: 31-Aug-1998
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
```

```
;             APPLICATION NUMBER: PCT/US97/03461
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: HOLLAND, DONALD R.
;             REGISTRATION NUMBER: 35,197
;             REFERENCE/DOCKET NUMBER: 976163
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (314) 727-5188
;             TELEFAX: (314) 727-6092
;   INFORMATION FOR SEQ ID NO: 12:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 585 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-08-981-739-12


  Query Match             10.1%;  Score 16;  DB 4;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db      24 GGCCGCTGCCTTCCAG 9


RESULT  40
US-08-931-858E-49/c
; Sequence 49, Application US/08931858E
; Patent No. 6222022
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON, EUGENE M
;    APPLICANT:  MILBRANDT, JEFFREY D
;    APPLICANT:  KOTZBAUER, PAUL T
;    APPLICANT:  LAMPE, PATRICIA A
;    APPLICANT:  KLEIN, ROBERT
;    APPLICANT:  DESAUVAGE, FRED
;    TITLE OF INVENTION:  PERSEPHIN AND RELATED GROWTH FACTOR
;    NUMBER OF SEQUENCES:  239
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MO
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/931,858E
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  971486
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  314-727-5188
;      TELEFAX:  314-727-6092
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1023 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-931-858E-49


  Query Match             10.1%;  Score 16;  DB 4;  Length 1023;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      38 ggccgctgccttccag 53
           ||||||||||||||||
Db     372 GGCCGCTGCCTTCCAG 357


RESULT  41
US-08-697-954-3/c
; Sequence 3, Application US/08697954
; Patent No. 6284535
;  GENERAL INFORMATION:
;    APPLICANT:  Role, Lorna W.
;    TITLE OF INVENTION:  SPLICE VARIANTS OF THE HEREGULIN GENE, nARIA, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/697,954
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  46839-A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-278-0400
;      TELEFAX:  212-391-0526
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1351 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-697-954-3


  Query Match             10.1%;  Score 16;  DB 4;  Length 1351;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     132 gaacctcactttcaga 147
           ||||||||||||||||
Db     552 GAACCTCACTTTCAGA 537


RESULT  42
US-08-428-926-1/c
; Sequence 1, Application US/08428926
; Patent No. 5667780
;  GENERAL INFORMATION:
;    APPLICANT:  Ho, Wei-Hsien
;    APPLICANT:  Osheroff, Phyllis L.
;    TITLE OF INVENTION:  SENSORY AND MOTOR NEURON DERIVED FACTOR (SMDF)
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
```

```
;       CITY:  South San Francisco
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  94080
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  patin (Genentech)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/428,926
;       FILING DATE:  25-APR-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/339517
;       FILING DATE:  14-NOV-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Lee, Wendy M.
;       REGISTRATION NUMBER:  00,000
;       REFERENCE/DOCKET NUMBER:  853D4
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415/225-1994
;       TELEFAX:  415/952-9881
;       TELEX:  910/371-7168
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1872 bases
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-428-926-1


  Query Match             10.1%;  Score 16;  DB 1;  Length 1872;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    132 gaacctcactttcaga 147
          ||||||||||||||||
Db    966 GAACCTCACTTTCAGA 951


RESULT  43
US-08-435-434-4/c
; Sequence 4, Application US/08435434
; Patent No. 5714385
;  GENERAL INFORMATION:
;    APPLICANT:  Mather, Jennie P.
;    APPLICANT:  Li, Ronghao
;    APPLICANT:  Chen, Jian
;    TITLE OF INVENTION:  ISOLATING AND CULTURING SCHWANN CELLS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/435,434
;      FILING DATE:  10-MAY-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lee, Wendy M.
;      REGISTRATION NUMBER:  00,000
;      REFERENCE/DOCKET NUMBER:  946-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-1994
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1872 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-435-434-4


  Query Match             10.1%;  Score 16;  DB 1;  Length 1872;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    132 gaacctcactttcaga 147
          ||||||||||||||||
Db    966 GAACCTCACTTTCAGA 951


RESULT  44
US-08-435-436-4/c
; Sequence 4, Application US/08435436
; Patent No. 5721139
;  GENERAL INFORMATION:
;    APPLICANT:  Mather, Jennie P.
;    APPLICANT:  Li, Ronghao
;    APPLICANT:  Chen, Jian
;    TITLE OF INVENTION:  ISOLATING AND CULTURING SCHWANN CELLS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/435,436
;      FILING DATE:  10-MAY-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lee, Wendy M.
;      REGISTRATION NUMBER:  00,000
;      REFERENCE/DOCKET NUMBER:  946-3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-1994
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1872 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-435-436-4


  Query Match             10.1%;  Score 16;  DB 1;  Length 1872;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy    132 gaacctcactttcaga 147
          ||||||||||||||||
Db    966 GAACCTCACTTTCAGA 951


RESULT  45
US-08-428-927-1/c
; Sequence 1, Application US/08428927
; Patent No. 5756456
;  GENERAL INFORMATION:
;    APPLICANT:  Ho, Wei-Hsien
;    APPLICANT:  Osheroff, Phyllis L.
;    TITLE OF INVENTION:  SENSORY AND MOTOR NEURON DERIVED FACTOR (SMDF)
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/428,927
;      FILING DATE:  25-APR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/339517
;      FILING DATE:  14-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lee, Wendy M.
;      REGISTRATION NUMBER:  00,000
;      REFERENCE/DOCKET NUMBER:  853D3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-1994
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1872 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-428-927-1


  Query Match             10.1%;  Score 16;  DB 1;  Length 1872;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    132 gaacctcactttcaga 147
          ||||||||||||||||
Db    966 GAACCTCACTTTCAGA 951


Search completed: September 20, 2002, 06:15:35
Job time: 11154 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 04:07:29 ; Search time 3900.56 Seconds
(without alignments)
550.181 Million cell updates/sec

Title: US-09-846-456-5
Perfect score: 159
Sequence: 1 ttaatgaccagccacgggcg.....ctttcagaagaagacaaaca 159

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	51	32.1	736	9	AU135588	AU135588 AU135588
	2	39	24.5	535	10	BG384217	BG384217 303216 MA
	3	24	15.1	292	10	Z44377	Z44377 HSC1ZB081 n
	4	20	12.6	661	10	BI391126	BI391126 pgpln.pk0
c	5	19	11.9	216	10	BF756949	BF756949 CM1-CT042
	6	19	11.9	454	12	AQ369174	AQ369174 HS_5032_A
c	7	19	11.9	470	9	AI353952	AI353952 zeh1204.s
c	8	19	11.9	493	12	AQ709464	AQ709464 HS_5459_A
c	9	19	11.9	618	9	AA495487	AA495487 fa09h12.r
c	10	19	11.9	643	9	AI497295	AI497295 fb63g04.y
	11	19	11.9	705	10	BI160520	BI160520 602864591
	12	19	11.9	730	9	AW116004	AW116004 fi06a11.x
	13	18	11.3	231	10	BF898523	BF898523 CM1-MT018
c	14	18	11.3	275	12	BH402163	BH402163 AG-ND-102
c	15	18	11.3	299	9	AA098714	AA098714 T4026 MVA
	16	18	11.3	303	10	BF903506	BF903506 IL2-MT017
	17	18	11.3	375	12	BH127062	BH127062 G-10g13.r
	18	18	11.3	397	10	BF911983	BF911983 IL2-UT007
c	19	18	11.3	529	12	BH396354	BH396354 AG-ND-161
c	20	18	11.3	532	12	AQ776697	AQ776697 HS_2148_B
	21	18	11.3	532	12	AQ779444	AQ779444 HS_2001_A
c	22	18	11.3	554	12	AZ236353	AZ236353 RPCI-23-7
c	23	18	11.3	558	12	AZ510640	AZ510640 1M0355G15
c	24	18	11.3	566	12	AZ510673	AZ510673 1M0355M15
c	25	18	11.3	599	9	AA720413	AA720413 ET2347 Tr
	26	18	11.3	631	9	AL652524	AL652524 AL652524
c	27	18	11.3	645	9	BB614720	BB614720 BB614720
	28	18	11.3	645	10	BI996668	BI996668 1031041B0
	29	18	11.3	647	12	CNS03C3I	AL237303 Tetraodon
	30	18	11.3	654	10	BG308188	BG308188 fm58c10.y
c	31	18	11.3	672	12	BH375609	BH375609 AG-ND-165
c	32	18	11.3	683	9	AW565995	AW565995 LG1_354_G
	33	18	11.3	689	9	AL648437	AL648437 AL648437
	34	18	11.3	693	9	AV868908	AV868908 AV868908
c	35	18	11.3	712	10	BI254723	BI254723 602978612
c	36	18	11.3	796	10	BI906710	BI906710 603064430
	37	18	11.3	809	12	BH181298	BH181298 018_I_20-
	38	18	11.3	809	12	CNS07MYE	AL618248 T3 end of
	39	18	11.3	855	12	BH475331	BH475331 BOGYG96TF
	40	18	11.3	887	10	BI464093	BI464093 603202870
	41	18	11.3	961	12	CNS02ELU	AL193899 Tetraodon
c	42	17	10.7	131	9	AW670238	AW670238 114155 MA
c	43	17	10.7	202	10	BF562309	BF562309 UI-R-BS0-
	44	17	10.7	218	9	BB183552	BB183552 BB183552
	45	17	10.7	249	9	AV061772	AV061772 AV061772

ALIGNMENTS

RESULT 1
AU135588
LOCUS AU135588 736 bp mRNA linear EST 24-OCT-2000
DEFINITION AU135588 PLACE1 Homo sapiens cDNA clone PLACE1002437 5', mRNA sequence.
ACCESSION AU135588
VERSION AU135588.1 GI:10996127
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 736)
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.
FEATURES Location/Qualifiers
source 1. .736
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="PLACE1002437"
/clone_lib="PLACE1"
/tissue_type="placenta"
/note="Vector: pME18SFL3"
BASE COUNT 163 a 199 c 199 g 170 t 5 others
ORIGIN

```
  Query Match             32.1%;  Score 51;  DB 9;  Length 736;
  Best Local Similarity   100.0%;  Pred. No. 6e-16;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    109 cagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 159
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    329 CAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAAGACAAACA 379


RESULT   2
BG384217
LOCUS       BG384217                 535 bp    mRNA    linear   EST 12-MAR-2001
DEFINITION  303216 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION   BG384217
VERSION     BG384217.1  GI:13308689
KEYWORDS    EST.
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 535)
  AUTHORS   Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
            Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
            and Keele,J.W.
  TITLE     Design and use of two pooled tissue normalized cDNA libraries for
            EST discovery in swine
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 90  row: G  column: 13
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
                     /clone_lib="MARC 1PIG"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                     Library made from pooled tissue from day 11, 13, 15, 20,
                     and 30 embryos."
BASE COUNT      121 a    159 c    136 g    119 t
ORIGIN


  Query Match             24.5%;  Score 39;  DB 10;  Length 535;
  Best Local Similarity   100.0%;  Pred. No. 9.9e-10;
  Matches   39;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    121 ctgctgtggaagaacctcactttcagaagaagacaaaca 159
          |||||||||||||||||||||||||||||||||||||||
Db    311 CTGCTGTGGAAGAACCTCACTTTCAGAAGAAGACAAACA 349


RESULT   3
Z44377
LOCUS       Z44377                   292 bp    mRNA    linear   EST 14-NOV-1994
DEFINITION  HSC1ZB081 normalized infant brain cDNA Homo sapiens cDNA clone
            c-1zb08, mRNA sequence.
ACCESSION   Z44377
VERSION     Z44377.1  GI:573506
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 292)
  AUTHORS   Auffray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C., Devignes
            ,M.D., Duprat,S., Houlgatte,R., Jumeau,M.N., Lamy,B., Lorenzo,F.,
            Mitchell,H., Mariage-Samson,R., Pietu,G., Pouliot,Y.,
            Sebastiani-Kabaktchis,C. and Tessier,A.
  TITLE     IMAGE: molecular integration of the analysis of the human genome
            and its expression
  JOURNAL   C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
  MEDLINE   95277534
COMMENT     Contact: Genethon
            Genexpress-Genethon
            Genethon Centre de recherche sur le Genome Humain
            1,rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
            Tel: 33169472800
            Fax: 33160778698
            Email: genexpress@genethon.fr
            Single read.
            Genexpress_library_idt: C; Genexpress_sequence_idt: y1c-1zb08
            Seq primer: (-21)M13_universal.
FEATURES             Location/Qualifiers
     source          1. .292
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="c-1zb08"
                     /clone_lib="normalized infant brain cDNA"
                     /sex="Female"
                     /tissue_type="total brain"
                     /dev_stage="3 months old"
                     /note="Organ: brain; Vector: lafmid BA; Site_1: HindIII;
                     Site_2: NotI; sex=Female; dev_stage=3 months old;
                     isolate=muscular atrophy patient;  tissue_type=total brain
                     ;  total mRNA was oligo-(dT) primed and directionally
                     cloned 5' -> 3' into the  HindIII -> NotI sites of the
                     lafmid BA vector. Clone library from B.Soares,  Psychiatry
                     Dept. Columbia University, USA.  Normalization_method:
                     Bento Soares, P.N.A.S in press"
BASE COUNT       50 a     87 c     96 g     56 t      3 others
ORIGIN


  Query Match             15.1%;  Score 24;  DB 10;  Length 292;
  Best Local Similarity   100.0%;  Pred. No. 0.056;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     32 agctctggccgctgccttccaggg 55
          |||||||||||||||||||||||
Db    236 AGCTCTGGCCGCTGCCTTCCAGGG 259


RESULT   4
BI391126
LOCUS       BI391126                 661 bp    mRNA    linear   EST 06-AUG-2001
DEFINITION  pgpln.pk003.g18 Normalized Chicken Pituitary/Hypothalamus/Pineal
            Library Gallus gallus cDNA clone pgpln.pk003.g18 5' similar to
            gi|13162285 ref|NP_076938.1| hypothetical protein MGC5242 [Homo
            sapiens] gb|AAH00168.1|AAH00168 (BC000168) Unknown (protein for
            MGC:5242) [Homo sapiens], mRNA sequence.
ACCESSION   BI391126
VERSION     BI391126.1  GI:15084408
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 661)
  AUTHORS   Porter,T.E. and Cogburn,L.A.
  TITLE     ESTs from Normalized Chicken Pituitary/Hypothalamus/Pineal cDNA
            library USDA/IFAFS Animal Genome Project
```

```
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Larry A. Cogburn
            University of Delaware
            Townsend Hall, Newark, DE 19717, USA
            Tel: 302-831-1335
            Fax: 302-831-2822
            Email: cogburn@udel.edu, www.chickest.udel.edu.
FEATURES             Location/Qualifiers
     source          1. .661
                     /organism="Gallus gallus"
                     /strain="Commercial broiler chicken"
                     /db_xref="taxon:9031"
                     /clone="pgp1n.pk003.g18"
                     /clone_lib="Normalized Chicken
                     Pituitary/Hypothalamus/Pineal Library"
                     /sex="Male and Female"
                     /tissue_type="Pituitary Gland/Hypothalamus/Pineal Gland"
                     /dev_stage="Embryonic (d12,d14,d19); post-hatch (w1,w3,w5
                     ,w7,w9)"
                     /lab_host="E. Coli EMDH10B"
                     /note="Vector: pCMVSPORT6; Library made from equivalent
                     pools of total RNA isolated from each  tissue at different
                     ages. Single pass sequencing from 5'-end"
BASE COUNT      155 a    164 c    225 g    117 t
ORIGIN


  Query Match             12.6%;  Score 20;  DB 10;  Length 661;
  Best Local Similarity   100.0%;  Pred. No. 8.7;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      69 cgctgggcgtgctggctgag 88
           ||||||||||||||||||||
Db      63 CGCTGGGCGTGCTGGCTGAG 82


RESULT   5
BF756949/c
LOCUS       BF756949                 216 bp    mRNA    linear   EST 12-JAN-2001
DEFINITION  CM1-CT0424-011100-522-e09 CT0424 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF756949
VERSION     BF756949.1  GI:12104849
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 216)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM1&t2=CM1-CT0424-
            011100-522-e09&t3=2000-11-01&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 216.
FEATURES             Location/Qualifiers
     source          1. .216
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="CT0424"
                     /dev_stage="Adult"
                     /note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No. 196
                     ,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
BASE COUNT       48 a     39 c     49 g     80 t
ORIGIN


  Query Match             11.9%;  Score 19;  DB 10;  Length 216;
  Best Local Similarity   100.0%;  Pred. No. 21;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     141 tttcagaagaagacaaaca 159
           ||||||||||||||||||||
Db     160 TTTCAGAAGAAGACAAACA 142


RESULT   6
AQ369174
LOCUS       AQ369174                 454 bp    DNA     linear   GSS 06-MAR-1999
DEFINITION  HS_5032_A2_F04_SP6E RPCI11 Human Male BAC Library Homo sapiens
            genomic clone Plate=608 Col=8 Row=K, DNA sequence.
ACCESSION   AQ369174
VERSION     AQ369174.1  GI:4338653
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 454)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 608  row: K  column: 8
            Seq primer: SP6
            Class: BAC ends
            High quality sequence stop: 454.
FEATURES             Location/Qualifiers
     source          1. .454
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=608 Col=8 Row=K"
                     /clone_lib="RPCI11 Human Male BAC Library"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /note="Vector: pBACe3.6; RPCI11 Human Male BAC Library"
BASE COUNT      144 a     82 c     90 g    136 t      2 others
ORIGIN


  Query Match             11.9%;  Score 19;  DB 12;  Length 454;
  Best Local Similarity   100.0%;  Pred. No. 26;
```

```
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      67 cacgctgggcgtgctggct 85
           ||||||||||||||||||||
Db     285 CACGCTGGGCGTGCTGGCT 303


RESULT   7
AI353952/c
LOCUS       AI353952                470 bp    mRNA    linear   EST 04-JAN-1999
DEFINITION  zeh1204.seq.F Zebrafish Embryonic Heart cDNA Library Danio rerio
            cDNA 5', mRNA sequence.
ACCESSION   AI353952
VERSION     AI353952.1  GI:4094105
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 470)
  AUTHORS   Ton,C., Mably,J.D., Dempsey,A.A., Hwang,D.M., Fishman,M.C. and Liew
            ,C.C.
  TITLE     Identification and Characterization of Expressed Sequence Tags from
            an Embryonic Zebrafish Heart cDNA Library
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Liew CC
            Brigham and Women's Hospital
            Harvard Medical School
            75 Francis St. Boston, MA 02115, USA
            Tel: 6177328915
            Fax: 6179750995
            Email: cliew@rics.bwh.harvard.edu
            PCR PRimers
            FORWARD: 5' GCCAAGCTCGAAATTAACCCTCACTAAAGGG 3'
            BACKWARD: 5' CCAGTGAATTGTAATACGACTCACTATAGGGCG 3'
            Seq primer: 5' GAAATTAACCCTCACTAAAGGG 3'.
FEATURES             Location/Qualifiers
     source          1. .470
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone_lib="Zebrafish Embryonic Heart cDNA Library"
                     /dev_stage="embryonic day 3 post-fertilization"
                     /lab_host="E.coli XL1-Blue mrF'"
                     /note="Organ: heart; Vector: Lambda ZAP Express; Site_1:
                     EcoRI; Site_2: XhoI; mRNA was purified from embryonic
                     zebrafish hearts (3 day post-fertilization). cDNA was
                     synthesized using a XhoI-Oligo dT adaptor-primer. EcoRI
                     adaptors were ligated, followed by digestion with XhoI,
                     for directional cloning into pre-digested lambda ZAP
                     Express vector. "
BASE COUNT      123 a    112 c    130 g    105 t
ORIGIN


  Query Match            11.9%;  Score 19;  DB 9;  Length 470;
  Best Local Similarity  100.0%;  Pred. No. 26;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     103 tggcctcagctgaggttgc 121
           ||||||||||||||||||||
Db     395 TGGCCTCAGCTGAGGTTGC 377


RESULT   8
AQ709464/c
LOCUS       AQ709464                493 bp    DNA     linear   GSS 07-JUL-1999
DEFINITION  HS_5459_A2_F04_T7A RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=1035 Col=8 Row=K, DNA sequence.
ACCESSION   AQ709464
VERSION     AQ709464.1  GI:5418890
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 493)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 1035  row: K  column: 8
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 493.
FEATURES             Location/Qualifiers
     source          1. .493
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=1035 Col=8 Row=K"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /sex="male"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
BASE COUNT      171 a     94 c     69 g    152 t      7 others
ORIGIN


  Query Match            11.9%;  Score 19;  DB 12;  Length 493;
  Best Local Similarity  100.0%;  Pred. No. 27;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     139 actttcagaagaagacaaa 157
           ||||||||||||||||||||
Db     478 ACTTTCAGAAGAAGACAAA 460


RESULT   9
AA495487/c
LOCUS       AA495487                618 bp    mRNA    linear   EST 27-JUN-1997
DEFINITION  fa09h12.r1 Zebrafish ICRFzfls Danio rerio cDNA clone 1108 5'
            similar to gb:J04973 UBIQUINOL-CYTOCHROME C REDUCTASE CORE PROTEIN
            2 PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION   AA495487
VERSION     AA495487.1  GI:2224979
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 618)
  AUTHORS   Clark,M., Lehrach,H., Appel,B., Eisen,J., Johnson,S., Marra,M.,
            Eddy,S., Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G.,
            Jost,S., Kucaba,T., Lacy,M., Le,N., Lennon,G., Martin,J., Moore,B.,
            Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
            ,T., Waterston,R. and Wilson,R.
```

```
  TITLE     WashU Zebrafish EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Steve Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Steve Johnson lab internal ID - P3_96 NOTE - For this library, the
            CLONE id field represents a position identifier on the original
            cDNA library preparation plate. cDNA Library Preparation: Matthew
            Clark. cDNA Library Arrayed by: Matthew Clark. DNA Sequencing by:
            Washington University Genome Sequencing Center Clone distribution:
            Genome Systems, St. Louis, and Max Planck Institut fuer Molekulare
            Genetik, Berlin Tel +49 30 84 13 1235
            Seq primer: T7 ET from Amersham
            High quality sequence stop: 427.
FEATURES             Location/Qualifiers
     source          1. .618
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone="1108"
                     /clone_lib="Zebrafish ICRFzfls"
                     /sex="mixed"
                     /tissue_type="pooled 26-somite embryos"
                     /lab_host="XL1-blue MRF"
                     /note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; 1st
                     strand cDNA was primed with a Not I - oligo(dT)15 primer
                     [5'pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT3'], on
                     mRNA from pooled 26 somite zebrafish embryos;
                     double-stranded cDNA was ligated to Sal I adaptors (BRL),
                     digested with Not I and cloned into the Not I and Sal I
                     sites of the pSPORT1 vector (BRL). Library was constructed
                     by Matthew Clark (Lehrach lab; ICRF, London and  Max
                     Planck Institut fuer Molekulare Genetik, Berlin) and was
                     not biochemically normalised.  70,000 clones from this
                     library were arrayed on high density filters and
                     subsequently screened by oligonucleotide hybridization
                     fingerprinting to identify unique or minimally redundant
                     clones for more intensive analysis."
BASE COUNT      150 a    161 c    168 g    139 t
ORIGIN
```

```
  Query Match             11.9%;  Score 19;  DB 9;  Length 618;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    103 tggcctcagctgaggttgc 121
          ||||||||||||||||||
Db    299 TGGCCTCAGCTGAGGTTGC 281
```

```
RESULT  10
AI497295/c
LOCUS       AI497295                643 bp    mRNA    linear   EST 07-JUN-2001
DEFINITION  fb63g04.y1 Zebrafish WashU MPIMG EST Danio rerio cDNA clone
            IMAGE:3716598 5' similar to SW:UCR2_HUMAN P22695
            UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX CORE PROTEIN 2 PRECURSOR
            ;, mRNA sequence.
ACCESSION   AI497295
VERSION     AI497295.1  GI:4398298
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 643)
  AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
            ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
            ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
            Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
            Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
            and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1998
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbrafish@watson.wustl.edu
            cDNA Library Preparation: Matthew Clark. cDNA Library Arrayed by:
            Matthew Clark. DNA Sequencing by: Washington University Genome
            Sequencing Center Clone distribution: Genome Systems, St. Louis,
            Missouri (web address: www.genomesystems.com) (email contact:
            info@genomesystems.com) and Research Genetics, Huntsville, Alabama
            (web address: www.resgen.com) (email contact: info@resgen.com) and
            RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address:
            www.rzpd.de)
            Possible reversed clone: similarity on wrong strand
            zebrafish identity (p-value greater than 1e-99) found to:
            gi|2224979|gb|AA495487|AA495487 fa09h12.r1 Zebrafish ICRFzfls Danio
            rerio cDNA
            Seq primer: T3 ET from Amersham
            High quality sequence stop: 479
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .643
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:3716598"
                     /clone_lib="Zebrafish WashU MPIMG EST"
                     /sex="mixed"
                     /tissue_type="26 somite embryos, adult livers, shield
                     stage embryos"
                     /lab_host="XL1-blue MRF"
                     /note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; 1st
                     strand cDNA was primed with a Not I - oligo(dT)15 primer
                     [5'pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT3'];
                     double-stranded cDNA was ligated to Sal I adaptors (BRL),
                     digested with Not I and cloned into the Not I and Sal I
                     sites of the pSPORT1 vector (BRL). Library was constructed
                     by Matthew Clark (Lehrach lab; ICRF, London and Max Planck
                     Institut fuer Molekulare Genetik,Berlin). cDNAs for EST
                     analysis were selected following oligonucleotide
                     hybridization fingerprinting of arrayed clones from
                     zebrafish late somitogenesis (26 ss), adult liver or
                     embryonic shield stage (5.6 h) libraries.  Fingerprint
                     data were used to computationally cluster cDNAs, and a
                     single cDNA from each cluster was chosen for sequencing.
                     In some cases multiple members of the same cluster were
                     sequenced to assess clustering parameters or single clones
                     were sequenced additional times to assess quality
                     control."
BASE COUNT      164 a    158 c    170 g    150 t      1 others
ORIGIN
```

```
  Query Match             11.9%;  Score 19;  DB 9;  Length 643;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    103 tggcctcagctgaggttgc 121
          ||||||||||||||||||
Db    395 TGGCCTCAGCTGAGGTTGC 377
```

```
RESULT  11
BI160520
LOCUS       BI160520                705 bp    mRNA    linear   EST 05-JUL-2001
DEFINITION  602864591F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:5018519 5',
            mRNA sequence.
ACCESSION   BI160520
VERSION     BI160520.1  GI:14620521
```

```
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 705)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1832  row: f  column: 24
            High quality sequence stop: 591.
FEATURES             Location/Qualifiers
     source          1. .705
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5018519"
                     /clone_lib="NIH_MGC_42"
                     /tissue_type="epithelioid carcinoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by Ling
                     Hong in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library. |"
BASE COUNT      152 a    209 c    224 g    120 t
ORIGIN


  Query Match             11.9%;  Score 19;  DB 10;  Length 705;
  Best Local Similarity   100.0%;  Pred. No. 30;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    110 agctgaggttgctgctgtg 128
          ||||||||||||||||||
Db    682 AGCTGAGGTTGCTGCTGTG 700


RESULT  12
AW116004
LOCUS       AW116004                730 bp    mRNA    linear   EST 20-OCT-1999
DEFINITION  fi06a11.x1 Sugano Kawakami zebrafish DRA Danio rerio cDNA clone
            2600348 3' similar to SW:UCR2_BOVIN P23004 UBIQUINOL-CYTOCHROME C
            REDUCTASE COMPLEX CORE PROTEIN 2 PRECURSOR ;, mRNA sequence.
ACCESSION   AW116004
VERSION     AW116004.1  GI:6082342
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 730)
  AUTHORS   Sugano,S.; Kawakami,K., Johnson,S., Li,F., Marra,M., Eddy,S.,
            Hillier,L., Clifton,S., Allen,M., Gibbons,M., Jost,S., Kucaba,T.,
            Martin,J., Pape,D., Steptoe,M., Underwood,K., Theising,B., Ritter
            ,E., Bowers,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: S.L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA
            Sequencing by: Washington University Genome Sequencing Center
            zebrafish identity (p-value greater than 1e-99) found to:
            gi|2224979|gb|AA495487|AA495487 fa09h12.r1 Zebrafish ICRFzfls Danio
            rerio cDNA
            Seq primer: T7 ET from Amersham
            High quality sequence stop: 450.
FEATURES             Location/Qualifiers
     source          1. .730
                     /organism="Danio rerio"
                     /strain="AB"
                     /db_xref="taxon:7955"
                     /clone="2600348"
                     /clone_lib="Sugano Kawakami zebrafish DRA"
                     /sex="mixed (one male and one female, including
                     unfertilized eggs)"
                     /dev_stage="adult"
                     /lab_host="DH10B (phage resistant)"
                     /note="Vector: pME18S-FL3; Site_1: DraIII (CACTGTGTG);
                     Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed
                     with an oligo(dT) primer  [ATGTGGCCTTTTTTTTTTTTTTTT];
                     double-stranded cDNA was ligated to a  DraIII adaptor
                     [TGTTGGCCTACTGG], digested and cloned into distinct DraIII
                     sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site
                     CACCATGTG). XhoI should be used to isolate the cDNA
                     insert.  Size selection was performed to exclude fragments
                     <1.5kb.  Library constructed by Dr. Sumio Sugano
                     (University of Tokyo Institute of Medical Science) and
                     kindly donated by Dr. Koichi Kawakami.  Custom primers for
                     sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end
                     primer CGACCTGCAGCTCGAGCACA.  "
BASE COUNT      195 a    167 c    168 g    198 t      2 others
ORIGIN


  Query Match             11.9%;  Score 19;  DB 9;  Length 730;
  Best Local Similarity   100.0%;  Pred. No. 30;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    103 tggcctcagctgaggttgc 121
          ||||||||||||||||||
Db    673 TGGCCTCAGCTGAGGTTGC 691


RESULT  13
BF898523
LOCUS       BF898523                231 bp    mRNA    linear   EST 18-JAN-2001
DEFINITION  CM1-MT0188-291100-611-d02 MT0188 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF898523
VERSION     BF898523.1  GI:12289982
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 231)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
```

```
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM1&t2=CM1-MT0188-
            291100-611-d02&t3=2000-11-29&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 5
            High quality sequence stop: 231.
FEATURES              Location/Qualifiers
     source           1. .231
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone_lib="MT0188"
                      /dev_stage="Adult"
                      /note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2:
                      SmaI; A mini-library was made by cloning products derived
                      from  ORESTES PCR (U.S. Letters Patent application No. 196
                      ,716 -  Ludwig Institute for Cancer Research) profiles
                      into  the pUC 18 vector. Reverse transcription of tissue
                      mRNA and cDNA amplification were performed under low
                      stringency conditions."
BASE COUNT       39 a     48 c     82 g     62 t
ORIGIN


  Query Match             11.3%;  Score 18;  DB 10;  Length 231;
  Best Local Similarity   100.0%;  Pred. No. 70;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     80 ctggctgagggaacatgg 97
          ||||||||||||||||||
Db    182 CTGGCTGAGGGAACATGG 199


RESULT  14
BH402163/c
LOCUS       BH402163                275 bp    DNA     linear   GSS 11-DEC-2001
DEFINITION  AG-ND-102E12.TF ND-TAM Anopheles gambiae genomic clone AG-ND-102E12
            , DNA sequence.
ACCESSION   BH402163
VERSION     BH402163.1  GI:17348379
KEYWORDS    GSS.
SOURCE      African malaria mosquito.
  ORGANISM  Anopheles gambiae
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
            ; Anopheles.
REFERENCE   1  (bases 1 to 275)
  AUTHORS   Shetty,J., Malek,J., Koo,H., Collins,F., Gardner,M. and Loftus,B.J.
  TITLE     Direct Submission of BAC-end sequences from Anopheles gambiae
  JOURNAL   Unpublished (2001)
COMMENT     Other_GSSs: AG-ND-102E12.TR
            Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            This clone is from an A. gambiae BAC library (ND-TAM) provided by
            F.H. Collins and sequenced by The Institute for Genomic Research
            (TIGR). The BAC library was generated from A. gambiae PEST strain
            DNA. All DNA was extracted from newly hatched first instar larvae
            to minimize the inclusion of DNA from microorganisms that inhabit
            the gut. The DNA is derived from mixed sexes of larvae. The BAC
            library was constructed at Texas A&M Universty BAC Center
            University, College Station, Texas 77843-2123, USA using a HindIII
            partial digest.
            Seq primer: M13 For
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .275
                      /organism="Anopheles gambiae"
                      /strain="PEST"
                      /db_xref="taxon:7165"
                      /clone="AG-ND-102E12"
                      /clone_lib="ND-TAM"
                      /note="Vector: pECBAC1; Site_1: HindIII"
BASE COUNT       73 a     47 c     73 g     82 t
ORIGIN


  Query Match             11.3%;  Score 18;  DB 12;  Length 275;
  Best Local Similarity   100.0%;  Pred. No. 74;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    139 actttcagaagaagacaa 156
          ||||||||||||||||||
Db    207 ACTTTCAGAAGAAGACAA 190


RESULT  15
AA098714/c
LOCUS       AA098714                299 bp    mRNA    linear   EST 28-OCT-1996
DEFINITION  T4026 MVAT4 bloodstream form of serodeme WRATat1.1 Trypanosoma
            brucei rhodesiense cDNA 5', mRNA sequence.
ACCESSION   AA098714
VERSION     AA098714.1  GI:1644736
KEYWORDS    EST.
SOURCE      Trypanosoma brucei rhodesiense.
  ORGANISM  Trypanosoma brucei rhodesiense
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 299)
  AUTHORS   Djikeng,A., Donelson,J.E. and Majiwa,P.A.O.
  TITLE     Generation of expressed sequence tags as physical landmarks in the
            genome of Trypanosoma brucei
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Majiwa PAO
            Molecular Biology Unit
            International Livestock Research Institute
            P.O. Box 30709, Nairobi, Kenya
            Tel: 254-2 630743
            Fax: 254-2 631499
            Email: p.majiwa@cgnet.com
            Seq primer: T3 primer.
FEATURES              Location/Qualifiers
     source           1. .299
                      /organism="Trypanosoma brucei rhodesiense"
                      /db_xref="taxon:31286"
                      /clone_lib="MVAT4 bloodstream form of serodeme WRATat1.1"
                      /note="Vector: Lambda ZAP II (Stratagene); Site_1: EcorI;
                      Site_2: XhoI; The mRNA was purified from a cloned
                      population of bloodstream trypanosomes reexpressing the
                      MVAT4 metacyclic variant surface glycoprotein (VSG). A
                      unidirectional oligo dT-primed EcoRI/XhoI cDNA library was
                      constructed in lambda ZAP II (Stratagene)."
BASE COUNT       97 a     60 c     94 g     48 t
ORIGIN


  Query Match             11.3%;  Score 18;  DB 9;  Length 299;
  Best Local Similarity   100.0%;  Pred. No. 76;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    111 gctgaggttgctgctgtg 128
          ||||||||||||||||||
Db    135 GCTGAGGTTGCTGCTGTG 118


RESULT  16
BF903506
```

```
LOCUS       BF903506                303 bp    mRNA    linear   EST 18-JAN-2001
DEFINITION  IL2-MT0179-181200-287-F06 MT0179 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF903506
VERSION     BF903506.1  GI:12295069
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 303)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL2&t2=IL2-MT0179-
            181200-287-F06&t3=2000-12-18&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 12
            High quality sequence stop: 283.
FEATURES             Location/Qualifiers
     source          1. .303
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="MT0179"
                     /dev_stage="Adult"
                     /note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No. 196
                     ,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
BASE COUNT       46 a     64 c    103 g     90 t
ORIGIN


  Query Match             11.3%;  Score 18;  DB 10;  Length 303;
  Best Local Similarity   100.0%;  Pred. No. 76;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      80 ctggctgagggaacatgg 97
           ||||||||||||||||||
Db     214 CTGGCTGAGGGAACATGG 231


RESULT  17
BH127062
LOCUS       BH127062                375 bp    DNA     linear   GSS 23-JUL-2001
DEFINITION  G-10g13.r Maize Random Small-insert Genomic Library Zea mays
            genomic clone G-10g13 both, DNA sequence.
ACCESSION   BH127062
VERSION     BH127062.1  GI:14994894
KEYWORDS    GSS.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 375)
  AUTHORS   Meyers,B.C., Tingey,S.V. and Morgante,M.
  TITLE     Abundance, distribution and transcriptional activity of repetitive
            elements in the maize genome
  JOURNAL   Genome Res. 11 (10), 1660-1676 (2001)
  MEDLINE   21475670
COMMENT     Contact: Morgante M
            Suite 200
            Dupont Genomics
            PO Box 6104, Newark, DE 19714-6104, USA
            Tel: 302 631 2638
            Fax: 302 631 2607
            Email: Michele.morgante@usa.dupont.com
            Sequences were trimmed to include only high quality bases; forward
            and reverse reads were assembled when significant overlaps were
            detected.
            Seq primer: M13reverse
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .375
                     /organism="Zea mays"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="G-10g13"
                     /clone_lib="Maize Random Small-insert Genomic Library"
                     /sex="hermaphrodite"
                     /tissue_type="leaf"
                     /cell_type="Young leaf"
                     /dev_stage="seedling"
                     /note="Vector: pCR-Script; Total genomic DNA was nebulized
                     ; ends were polished with Pfu polymerase and the fragments
                     cloned into pCR-Script."
BASE COUNT       99 a     74 c     84 g    111 t      7 others
ORIGIN


  Query Match             11.3%;  Score 18;  DB 12;  Length 375;
  Best Local Similarity   100.0%;  Pred. No. 81;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      84 ctgagggaacatggcatg 101
           ||||||||||||||||||
Db       7 CTGAGGGAACATGGCATG 24


RESULT  18
BF911983
LOCUS       BF911983                397 bp    mRNA    linear   EST 18-JAN-2001
DEFINITION  IL2-UT0073-121100-231-A10 UT0073 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF911983
VERSION     BF911983.1  GI:12303441
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 397)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
```

```
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL2&t2=IL2-UT0073-
            121100-231-A10&t3=2000-11-12&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 372.
FEATURES             Location/Qualifiers
     source          1. .397
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="UT0073"
                     /dev_stage="Adult"
                     /note="Organ: uterus_tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       75 a    112 c     87 g    123 t
ORIGIN


  Query Match             11.3%;  Score 18;  DB 10;  Length 397;
  Best Local Similarity   100.0%;  Pred. No. 83;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     120 gctgctgtggaagaacct 137
           ||||||||||||||||||
Db      57 GCTGCTGTGGAAGAACCT 74


RESULT  19
BH396354/c
LOCUS       BH396354                529 bp    DNA     linear   GSS 11-DEC-2001
DEFINITION  AG-ND-161J6.TF ND-TAM Anopheles gambiae genomic clone AG-ND-161J6,
            DNA sequence.
ACCESSION   BH396354
VERSION     BH396354.1  GI:17342495
KEYWORDS    GSS.
SOURCE      African malaria mosquito.
  ORGANISM  Anopheles gambiae
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
            ; Anopheles.
REFERENCE   1  (bases 1 to 529)
  AUTHORS   Shetty,J., Malek,J., Koo,H., Collins,F., Gardner,M. and Loftus,B.J.
  TITLE     Direct Submission of BAC-end sequences from Anopheles gambiae
  JOURNAL   Unpublished (2001)
COMMENT     Other_GSSs: AG-ND-161J6.TR
            Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            This clone is from an A. gambiae BAC library (ND-TAM) provided by
            F.H. Collins and sequenced by The Institute for Genomic Research
            (TIGR). The BAC library was generated from A. gambiae PEST strain
            DNA. All DNA was extracted from newly hatched first instar larvae
            to minimize the inclusion of DNA from microorganisms that inhabit
            the gut. The DNA is derived from mixed sexes of larvae. The BAC
            library was constructed at Texas A&M Universty BAC Center
            University, College Station, Texas 77843-2123, USA using a HindIII
            partial digest.
            Seq primer: M13 For
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .529
                     /organism="Anopheles gambiae"
                     /strain="PEST"
                     /db_xref="taxon:7165"
                     /clone="AG-ND-161J6"
                     /clone_lib="ND-TAM"
                     /note="Vector: pECBAC1; Site_1: HindIII"
BASE COUNT      122 a    121 c    151 g    135 t
ORIGIN


  Query Match             11.3%;  Score 18;  DB 12;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 90;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     139 actttcagaagaagacaa 156
           ||||||||||||||||||
Db     206 ACTTTCAGAAGAAGACAA 189


RESULT  20
AQ776697/c
LOCUS       AQ776697                532 bp    DNA     linear   GSS 29-JUL-1999
DEFINITION  HS_2148_B2_H11_T7C CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2148 Col=22 Row=P, DNA sequence.
ACCESSION   AQ776697
VERSION     AQ776697.1  GI:5656425
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 532)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 2148  row: P  column: 22
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 532.
FEATURES             Location/Qualifiers
     source          1. .532
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=2148 Col=22 Row=P"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /sex="male"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
BASE COUNT      200 a     94 c     83 g    141 t     14 others
ORIGIN


  Query Match             11.3%;  Score 18;  DB 12;  Length 532;
  Best Local Similarity   100.0%;  Pred. No. 90;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      86 gagggaacatggcatgtt 103
           ||||||||||||||||||
Db     226 GAGGGAACATGGCATGTT 209
```

```
RESULT  21
AQ779444
LOCUS       AQ779444                 532 bp    DNA     linear   GSS 02-AUG-1999
DEFINITION  HS_2001_A2_F04_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2001 Col=8 Row=K, DNA sequence.
ACCESSION   AQ779444
VERSION     AQ779444.1  GI:5682404
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 532)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 2001  row: K  column: 8
            Seq primer: M13 Reverse
            Class: BAC ends
            High quality sequence stop: 532.
FEATURES             Location/Qualifiers
     source          1. .532
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=2001 Col=8 Row=K"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /sex="male"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
BASE COUNT      155 a     86 c    156 g    126 t      9 others
ORIGIN


  Query Match             11.3%;  Score 18;  DB 12;  Length 532;
  Best Local Similarity   100.0%;  Pred. No. 90;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    118 ttgctgctgtggaagaac 135
          ||||||||||||||||||
Db    372 TTGCTGCTGTGGAAGAAC 389


RESULT  22
AZ236353/c
LOCUS       AZ236353                 554 bp    DNA     linear   GSS 14-JUN-2000
DEFINITION  RPCI-23-70H19.TV RPCI-23 Mus musculus genomic clone RPCI-23-70H19,
            DNA sequence.
ACCESSION   AZ236353
VERSION     AZ236353.1  GI:8544399
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 554)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
            ,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-70H19.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 70  row: H  column: 19
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .554
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-70H19"
                     /clone_lib="RPCI-23"
                     /sex="Female"
                     /lab_host="DH10B"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and  partially digested
                     with a combination of EcoRI and EcoRI Methylase.   Size
                     selected DNA was cloned into the pBACe3.6 vector at the
                     EcoRI  sites.  The ligation products were transformed into
                     DH10B  electrocompetent cells (BRL Life Technologies).
BASE COUNT      141 a    129 c    131 g    153 t
ORIGIN


  Query Match             11.3%;  Score 18;  DB 12;  Length 554;
  Best Local Similarity   100.0%;  Pred. No. 92;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    109 cagctgaggttgctgctg 126
          ||||||||||||||||||
Db    407 CAGCTGAGGTTGCTGCTG 390


RESULT  23
AZ510640/c
LOCUS       AZ510640                 558 bp    DNA     linear   GSS 05-OCT-2000
DEFINITION  1M0355G15F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0355G15 F, DNA sequence.
ACCESSION   AZ510640
VERSION     AZ510640.1  GI:10691956
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 558)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
```

```
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0355  row: G  column: 15
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 558.
FEATURES             Location/Qualifiers
     source          1. .558
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0355G15"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT      189 a    110 c     86 g    173 t
ORIGIN


  Query Match             11.3%;  Score 18;  DB 12;  Length 558;
  Best Local Similarity   100.0%;  Pred. No. 92;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     118 ttgctgctgtggaagaac 135
           ||||||||||||||||||
Db     255 TTGCTGCTGTGGAAGAAC 238


RESULT  24
AZ510673/c
LOCUS       AZ510673                 566 bp    DNA     linear   GSS 05-OCT-2000
DEFINITION  1M0355M15F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0355M15 F, DNA sequence.
ACCESSION   AZ510673
VERSION     AZ510673.1  GI:10691989
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 566)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0355  row: M  column: 15
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 566.
FEATURES             Location/Qualifiers
     source          1. .566
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0355M15"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT      193 a    103 c     85 g    185 t
ORIGIN


  Query Match             11.3%;  Score 18;  DB 12;  Length 566;
  Best Local Similarity   100.0%;  Pred. No. 92;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     118 ttgctgctgtggaagaac 135
           ||||||||||||||||||
Db     233 TTGCTGCTGTGGAAGAAC 216


RESULT  25
AA720413/c
LOCUS       AA720413                 599 bp    mRNA    linear   EST 30-DEC-1997
DEFINITION  ET2347 Trypanosoma brucei rhodesiense ZAP II library Trypanosoma
            brucei rhodesiense cDNA 5', mRNA sequence.
ACCESSION   AA720413
VERSION     AA720413.1  GI:2734023
KEYWORDS    EST.
SOURCE      Trypanosoma brucei rhodesiense.
  ORGANISM  Trypanosoma brucei rhodesiense
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 599)
  AUTHORS   Ullu,E. and Tschudi,C.
  TITLE     Expressed sequence tags from procyclic Trypanosoma brucei
            rhodesiense cDNA clones
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Ullu E
            Department of Internal Medicine, Section of Infectious Diseases
            Yale University School of Medicine
            P.O. Box 208022, 333 Cedar Street, New Haven, CT 06520-8022, USA
            Fax: 203 785 3864
            Email: elisabetta.ullu@yale.edu
            Seq primer: SK.
```

```
FEATURES             Location/Qualifiers
     source          1. .599
                     /organism="Trypanosoma brucei rhodesiense"
                     /strain="YTat 1.1"
                     /db_xref="taxon:31286"
                     /clone_lib="Trypanosoma brucei rhodesiense ZAP II library"
                     /dev_stage="Insect form"
                     /note="Vector: Lambda ZAP II; Site_1: Eco RI; Site_2: Xho
                     I; A unidirectional oligo dT-primed cDNA library was
                     constructed in lambda ZAP II. Clones were selected using
                     the criteria of low reactivity with a total cDNA probe."
BASE COUNT      167 a    136 c    131 g    127 t     38 others
ORIGIN


  Query Match             11.3%;  Score 18;  DB 9;  Length 599;
  Best Local Similarity   100.0%;  Pred. No. 94;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    111 gctgaggttgctgctgtg 128
          ||||||||||||||||||
Db    293 GCTGAGGTTGCTGCTGTG 276


RESULT  26
AL652524
LOCUS       AL652524                631 bp    mRNA    linear   EST 13-DEC-2001
DEFINITION  AL652524 XGC-gastrula Silurana tropicalis cDNA clone TGas028h10 5',
            mRNA sequence.
ACCESSION   AL652524
VERSION     AL652524.1  GI:17663079
KEYWORDS    EST.
SOURCE      western clawed frog.
  ORGANISM  Silurana tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Silurana.
REFERENCE   1  (bases 1 to 631)
  AUTHORS   Huckle,E., Taylor,R., Ashurst,J.L., Zorn,A.M. and Rogers,J.
  TITLE     Sanger Xenopus tropicalis EST project 2001 (10_2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Huckle E
            Sanger Centre
            Hinxton, Cambridgeshire, CB10 1SA, UK
            Email: trop@sanger.ac.uk
            Sanger Xenopus tropicalis EST project 2001
            TROPICALIS_SEQUENCE_ID: TGas028h10.sp6
            Sequencing primer: SP6
            This sequence is from a Xenopus Gene Collection (XGC) library
            constructed by Aaron M. Zorn.
FEATURES             Location/Qualifiers
     source          1. .631
                     /organism="Silurana tropicalis"
                     /db_xref="taxon:8364"
                     /clone="TGas028h10"
                     /clone_lib="XGC-gastrula"
                     /dev_stage="gastrula (stages 10.5-13 mixed)"
                     /lab_host="Escherichia coli XL1-blue"
                     /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
                     was oligo dT primed from 5ug of poly A+ RNA from stages
                     10-13 gastrulae.  EcoRI-NotI cut cDNA was then ligated
                     into pCS107 with EcoRI at the 5' end and NotI at the 3'
                     end."
BASE COUNT      189 a    171 c    161 g    110 t
ORIGIN


  Query Match             11.3%;  Score 18;  DB 9;  Length 631;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     61 gagccacacgctgggcgt 78
          ||||||||||||||||||
Db     78 GAGCCACACGCTGGGCGT 95


RESULT  27
BB614720/c
LOCUS       BB614720                645 bp    mRNA    linear   EST 26-OCT-2001
DEFINITION  BB614720 RIKEN full-length enriched, adult male testis Mus musculus
            cDNA clone 4921533D20 5', mRNA sequence.
ACCESSION   BB614720
VERSION     BB614720.1  GI:16455076
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 645)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
            , Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
            ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
            Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
            ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .645
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="4921533D20"
                     /clone_lib="RIKEN full-length enriched, adult male testis"
                     /sex="male"
                     /tissue_type="testis"
                     /dev_stage="adult"
                     /lab_host="SOLR"
                     /note="Site_1: XhoI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
```

Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGAGCGGCCGCAATTAATTCTCGAGTTAATTAAATTAATCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. "

BASE COUNT 175 a 185 c 172 g 113 t
ORIGIN

Query Match 11.3%; Score 18; DB 9; Length 645;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 109 cagctgaggttgctgctg 126
||||||||||||||||||
Db 454 CAGCTGAGGTTGCTGCTG 437

RESULT 28
BI996668
LOCUS BI996668 645 bp mRNA linear EST 25-OCT-2001
DEFINITION 1031041B01.y2 C. reinhardtii CC-1690, Stress II (normalized), Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION BI996668
VERSION BI996668.1 GI:16431442
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadaceae; Chlamydomonas.
REFERENCE 1 (bases 1 to 645)
AUTHORS Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C., Lefebvre,P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants. Project: 1031
JOURNAL Unpublished (2001)
COMMENT Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.
FEATURES Location/Qualifiers
source 1. .645
/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized), Lambda Zap II"
/note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2: XhoI; Stress condition II library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (NH4+ - containing) and shifted to TAP - NO3- (24hrs); H2 production conditions (0, 12hr, 24hr) see Melis et al.,(2000) Plant Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP + sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr). PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoRI (3') sites. pBluescript II SK- plasmids were excised from the lambda ZAP clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806."

BASE COUNT 137 a 196 c 242 g 70 t
ORIGIN

Query Match 11.3%; Score 18; DB 10; Length 645;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 31 cagctctggccgctgcct 48
||||||||||||||||||
Db 258 CAGCTCTGGCCGCTGCCT 275

RESULT 29
CNS03C3I
LOCUS CNS03C3I 647 bp DNA linear GSS 15-MAY-2000
DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone 013H01 of library G from Tetraodon nigroviridis, genomic survey sequence.
ACCESSION AL237303
VERSION AL237303.1 GI:7896438
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 647)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
TITLE Charaterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 647)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.
TITLE Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 647)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) to the EMBL/GenBank/DDBJ databases
COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon.
FEATURES Location/Qualifiers
source 1. .647
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="013H01"
/clone_lib="G"
/note="Genoscope sequence ID : C0BG013CD01LP1-end : T7"
BASE COUNT 109 a 253 c 142 g 119 t 24 others
ORIGIN

Query Match 11.3%; Score 18; DB 12; Length 647;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 128 ggaagaacctcactttca 145
||||||||||||||||||
Db 505 GGAAGAACCTCACTTTCA 522

RESULT 30
BG308188
LOCUS BG308188 654 bp mRNA linear EST 13-FEB-2002

```
DEFINITION  fm58c10.y1 Zebrafish adult retina cDNA Danio rerio cDNA clone
            4199443 5' similar to SW:GBT1_BOVIN P04695 GUANINE
            NUCLEOTIDE-BINDING PROTEIN G(T), ALPHA-1 SUBUNIT ;, mRNA sequence.
ACCESSION   BG308188
VERSION     BG308188.1  GI:13105715
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 654)
  AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
            ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
            ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
            Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
            Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
            and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1998
  JOURNAL   Unpublished (1998)
COMMENT     Other_ESTs: fm58c10.x1
            Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbrafish@watson.wustl.edu
            Library constructed by: Chandra Tucker and Gregory Niemi DNA
            Sequencing by: Washington University Genome Sequencing Center Clone
            distribution: RessourcenZentrumPrimarDatenbank, Berlin, Germany
            (web address: www.rzpd.de)
            Seq primer: T3 ET from Amersham
            High quality sequence stop: 410.
FEATURES             Location/Qualifiers
     source          1. .654
                     /organism="Danio rerio"
                     /strain="wild-type"
                     /db_xref="taxon:7955"
                     /clone="4199443"
                     /clone_lib="Zebrafish adult retina cDNA"
                     /sex="mixed"
                     /dev_stage="1-2 years"
                     /lab_host="E.Coli XL1-Blue MRF' (XL1-Blue MRF')"
                     /note="Vector: Lambda ZAP II (pBluescript SK-); Site_1:
                     EcoRI; Site_2: SalI; This Zebrafish library was
                     constructed by Dr. Susan E. Brockerhoff (email:
                     sbrocker@u.washington.edu) RZPD library number: 760"
BASE COUNT      189 a    150 c    170 g    145 t
ORIGIN


  Query Match             11.3%;  Score 18;  DB 10;  Length 654;
  Best Local Similarity   100.0%;  Pred. No. 96;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     130 aagaacctcactttcaga 147
           ||||||||||||||||||
Db     635 AAGAACCTCACTTTCAGA 652


RESULT  31
BH375609/c
LOCUS       BH375609                672 bp    DNA     linear   GSS 10-DEC-2001
DEFINITION  AG-ND-165M6.TF ND-TAM Anopheles gambiae genomic clone AG-ND-165M6,
            DNA sequence.
ACCESSION   BH375609
VERSION     BH375609.1  GI:17321751
KEYWORDS    GSS.
SOURCE      African malaria mosquito.
  ORGANISM  Anopheles gambiae
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
            ; Anopheles.
REFERENCE   1  (bases 1 to 672)
  AUTHORS   Shetty,J., Malek,J., Koo,H., Collins,F., Gardner,M. and Loftus,B.J.
  TITLE     Direct Submission of BAC-end sequences from Anopheles gambiae
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            This clone is from an A. gambiae BAC library (ND-TAM) provided by
            F.H. Collins and sequenced by The Institute for Genomic Research
            (TIGR). The BAC library was generated from A. gambiae PEST strain
            DNA. All DNA was extracted from newly hatched first instar larvae
            to minimize the inclusion of DNA from microorganisms that inhabit
            the gut. The DNA is derived from mixed sexes of larvae. The BAC
            library was constructed at Texas A&M Universty BAC Center
            University, College Station, Texas 77843-2123, USA using a HindIII
            partial digest.
            Seq primer: M13 For
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .672
                     /organism="Anopheles gambiae"
                     /strain="PEST"
                     /db_xref="taxon:7165"
                     /clone="AG-ND-165M6"
                     /clone_lib="ND-TAM"
                     /note="Vector: pECBAC1; Site_1: HindIII"
BASE COUNT      176 a    155 c    160 g    181 t
ORIGIN


  Query Match             11.3%;  Score 18;  DB 12;  Length 672;
  Best Local Similarity   100.0%;  Pred. No. 97;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ttaatgaccagccacggg 18
           ||||||||||||||||||
Db     612 TTAATGACCAGCCACGGG 595


RESULT  32
AW565995/c
LOCUS       AW565995                683 bp    mRNA    linear   EST 19-JUL-2000
DEFINITION  LG1_354_G11.g1_A002 Light Grown 1 (LG1) Sorghum bicolor cDNA, mRNA
            sequence.
ACCESSION   AW565995
VERSION     AW565995.1  GI:7219873
KEYWORDS    EST.
SOURCE      sorghum.
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 683)
  AUTHORS   Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C. and Pratt,L.H.
  TITLE     An EST database from Sorghum: light-grown seedlings
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Cordonnier-Pratt MM
            Department of Botany
            The University of Georgia
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 542 1805
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20.
            Seq primer: T7
            High quality sequence start: 4
            High quality sequence stop: 667
```

```
           POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .683
                     /organism="Sorghum bicolor"
                     /db_xref="taxon:4558"
                     /clone_lib="Light Grown 1 (LG1)"
                     /note="Organ: 10- to 14-day-old light-grown (greenhouse)
                     seedlings; Vector: Lambda Zap; Site_1: XhoI; Site_2: EcoRI
                     ; The library was made from poly-A RNA in the cloning
                     vector lambda ZAP II.  Clones to be sequenced were
                     prepared by mass excision."
BASE COUNT      206 a    123 c    164 g    189 t      1 others
ORIGIN


  Query Match             11.3%;  Score 18;  DB 9;  Length 683;
  Best Local Similarity   100.0%;  Pred. No. 98;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    130 aagaacctcactttcaga 147
          ||||||||||||||||||
Db    553 AAGAACCTCACTTTCAGA 536


RESULT  33
AL648437
LOCUS       AL648437                689 bp    mRNA    linear   EST 13-DEC-2001
DEFINITION  AL648437 XGC-gastrula Silurana tropicalis cDNA clone TGas033i21 5',
            mRNA sequence.
ACCESSION   AL648437
VERSION     AL648437.1  GI:17657152
KEYWORDS    EST.
SOURCE      western clawed frog.
  ORGANISM  Silurana tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Silurana.
REFERENCE   1  (bases 1 to 689)
  AUTHORS   Huckle,E., Taylor,R., Ashurst,J.L., Zorn,A.M. and Rogers,J.
  TITLE     Sanger Xenopus tropicalis EST project 2001 (10_2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Huckle E
            Sanger Centre
            Hinxton, Cambridgeshire, CB10 1SA, UK
            Email: trop@sanger.ac.uk
            Sanger Xenopus tropicalis EST project 2001
            TROPICALIS_SEQUENCE_ID: TGas033i21.sp6
            Sequencing primer: SP6
            This sequence is from a Xenopus Gene Collection (XGC) library
            constructed by Aaron M. Zorn.
FEATURES             Location/Qualifiers
     source          1. .689
                     /organism="Silurana tropicalis"
                     /db_xref="taxon:8364"
                     /clone="TGas033i21"
                     /clone_lib="XGC-gastrula"
                     /dev_stage="gastrula (stages 10.5-13 mixed)"
                     /lab_host="Escherichia coli XL1-blue"
                     /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
                     was oligo dT primed from 5ug of poly A+ RNA from stages
                     10-13 gastrulae.  EcoRI-NotI cut cDNA was then ligated
                     into pCS107 with EcoRI at the 5' end and NotI at the 3'
                     end."
BASE COUNT      184 a    193 c    185 g    127 t
ORIGIN


  Query Match             11.3%;  Score 18;  DB 9;  Length 689;
  Best Local Similarity   100.0%;  Pred. No. 98;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     61 gagccacacgctgggcgt 78
          ||||||||||||||||||
Db    173 GAGCCACACGCTGGGCGT 190


RESULT  34
AV868908
LOCUS       AV868908                693 bp    mRNA    linear   EST 08-NOV-2001
DEFINITION  AV868908 Nori Satoh unpublished cDNA library, egg Ciona
            intestinalis cDNA clone rcieg35d09 3', mRNA sequence.
ACCESSION   AV868908
VERSION     AV868908.1  GI:16856432
KEYWORDS    EST.
SOURCE      Ciona intestinalis.
  ORGANISM  Ciona intestinalis
            Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
            Phlebobranchia; Cionidae; Ciona.
REFERENCE   1  (bases 1 to 693)
  AUTHORS   Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.
  TITLE     Expressed genes in Ciona intestinalis
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Nori Satoh
            Department of Zoology
            Kyoto University
            Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
            Tel: 81-75-753-4081
            Fax: 81-75-705-1113
            Email: satoh@ascidian.zool.kyoto-u.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .693
                     /organism="Ciona intestinalis"
                     /db_xref="taxon:7719"
                     /clone="rcieg35d09"
                     /clone_lib="Nori Satoh unpublished cDNA library, egg"
                     /tissue_type="whole animal"
                     /dev_stage="egg"
BASE COUNT      195 a    142 c    120 g    235 t      1 others
ORIGIN


  Query Match             11.3%;  Score 18;  DB 9;  Length 693;
  Best Local Similarity   100.0%;  Pred. No. 98;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    135 cctcactttcagaagaag 152
          ||||||||||||||||||
Db    247 CCTCACTTTCAGAAGAAG 264


RESULT  35
BI254723/c
LOCUS       BI254723                712 bp    mRNA    linear   EST 17-JUL-2001
DEFINITION  602978612F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5123378 5',
            mRNA sequence.
ACCESSION   BI254723
VERSION     BI254723.1  GI:14807426
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 712)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Incyte Genomics, Inc.
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11301  row: h  column: 03
```

```
            High quality sequence stop: 710.
FEATURES             Location/Qualifiers
     source          1. .712
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5123378"
                     /clone_lib="NIH_MGC_12"
                     /tissue_type="cervical carcinoma cell line"
                     /lab_host="DH10B"
                     /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.4 kb. Library prepared by Life
                     Technologies."
BASE COUNT      197 a    176 c    176 g    163 t
ORIGIN


  Query Match             11.3%;  Score 18;  DB 10;  Length 712;
  Best Local Similarity   100.0%;  Pred. No. 99;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      79 gctggctgagggaacatg 96
           ||||||||||||||||||
Db     263 GCTGGCTGAGGGAACATG 246


RESULT  36
BI906710/c
LOCUS       BI906710                796 bp    mRNA    linear   EST 16-OCT-2001
DEFINITION  603064430F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5213695 5',
            mRNA sequence.
ACCESSION   BI906710
VERSION     BI906710.1  GI:16169467
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 796)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11536  row: k  column: 08
            High quality sequence stop: 787.
FEATURES             Location/Qualifiers
     source          1. .796
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5213695"
                     /clone_lib="NIH_MGC_118"
                     /tissue_type="leukocyte"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
                     (destroyed); RNA source leukocytes from anonymous pool of
                     non-activated  adult donors.  Library is oligo-dT primed
                     and directionally cloned  (EcoRV site is destroyed upon
                     cloning).  Average insert size 1.7 kb, insert size range
                     1.2-3.3 kb.  Library is normalized and enriched for
                     full-length clones and was constructed by C. Gruber
                     (Invitrogen).  Research Genetics tracking code 027. Note:
                     this is a NIH_MGC Library."
BASE COUNT      209 a    206 c    192 g    189 t
ORIGIN


  Query Match             11.3%;  Score 18;  DB 10;  Length 796;
  Best Local Similarity   100.0%;  Pred. No. 1e+02;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     131 agaacctcactttcagaa 148
           ||||||||||||||||||
Db      20 AGAACCTCACTTTCAGAA 3


RESULT  37
BH181298
LOCUS       BH181298                809 bp    DNA     linear   GSS 29-OCT-2001
DEFINITION  018_I_20-rev SmBAC1 Schistosoma mansoni genomic clone 018I20 5',
            DNA sequence.
ACCESSION   BH181298
VERSION     BH181298.1  GI:16284183
KEYWORDS    GSS.
SOURCE      Schistosoma mansoni.
  ORGANISM  Schistosoma mansoni
            Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
            Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.
REFERENCE   1  (bases 1 to 809)
  AUTHORS   Le Paslier,M.-C., Pierce,R.J., Merlin,F., Hirai,H., Wu,W., Williams
            ,D.L., Johnston,D., LoVerde,P.T. and Le Paslier,D.
  TITLE     Construction and characterization of a Schistosoma mansoni
            bacterial artificial chromosome library
  JOURNAL   Genomics 65 (2), 87-94 (2000)
  MEDLINE   20247747
COMMENT     Other_GSSs: 018_I_20-21
            Contact: Pierce RJ
            INSERM U 167
            Institut Pasteur de Lille
            1 rue du Professeur A. Calmette, 59019-Lille, France
            Tel: (33) (0)3 20877783
            Fax: (33) (0)3 20877888
            Email: Raymond.Pierce@pasteur-lille.fr
            CNS sequencing ID=DG0AA018BE10BP1 Bases 167-538 have 84% identity
            to S.mansoni EST AW061395.1 from base 25-393.
            Plate: 018  row: I  column: 20
            Seq primer: M13 reverse primer
            Class: BAC ends
            High quality sequence stop: 809.
FEATURES             Location/Qualifiers
     source          1. .809
                     /organism="Schistosoma mansoni"
                     /strain="Puerto-Rican"
                     /db_xref="taxon:6183"
                     /clone="018I20"
                     /clone_lib="SmBAC1"
                     /sex="mixed"
                     /dev_stage="cercariae"
                     /lab_host="Biomphalaria glabrata"
                     /note="Vector: pBeloBAC 11; Site_1: Hind III; Partially
                     Hind III digested and size-selected S. mansoni  cercarial
                     DNA was ligated into Hind III digested pBeloBAC 11 vector
                     and used to transform E. coli DH10B. The complete library
                     contains 23808 clones from 4 independent
                     sizing-ligation-transformations.  Average insert size
                     ranges from 70-127 kb and genome coverage is 7.9-fold."
BASE COUNT      313 a    155 c    109 g    229 t      3 others
ORIGIN


  Query Match             11.3%;  Score 18;  DB 12;  Length 809;
  Best Local Similarity   100.0%;  Pred. No. 1e+02;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     142 ttcagaagaagacaaaca 159
           ||||||||||||||||||
Db     536 TTCAGAAGAAGACAAACA 553
```